

Spring 2007

Analysis of Georgia Elementary Principals' Perceptions of High-Stakes Testing

Sabrina Vaughn Calhoun

Follow this and additional works at: <https://digitalcommons.georgiasouthern.edu/etd>

Recommended Citation

Calhoun, Sabrina Vaughn, "Analysis of Georgia Elementary Principals' Perceptions of High-Stakes Testing" (2007). *Electronic Theses and Dissertations*. 205.

<https://digitalcommons.georgiasouthern.edu/etd/205>

This dissertation (open access) is brought to you for free and open access by the Graduate Studies, Jack N. Averitt College of at Digital Commons@Georgia Southern. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of Digital Commons@Georgia Southern. For more information, please contact digitalcommons@georgiasouthern.edu.

AN ANALYSIS OF GEORGIA ELEMENTARY PRINCIPALS' PERCEPTIONS OF
HIGH-STAKES TESTING

by

SABRINA VAUGHN CALHOUN

(Under the Direction of Michael D. Richardson)

ABSTRACT

The study explored the perceptions of the Georgia elementary principals as high-stakes testing is utilized in the theme of accountability. The study examined 335 principals' personal and professional demographics and perceptions of the implementation of high-stakes testing.

The study employed a descriptive, survey approach to address the research questions. A self-designed survey questionnaire was developed to explore principals' perception of high-stakes testing, and included both a qualitative and quantitative orientation.

Findings indicated that the majority of the 335 Georgia elementary principals who responded to the survey were 46-55 year old females who worked in suburban areas of the state. They typically held the Education Specialist degree, have an average of two years of experience as principal, and made AYP for the 2005-06 school year.

Respondents believed that they could use high-stakes testing results to improve student achievement. At the same time, the principals expressed concern that factors beyond the control of the principal influenced student achievement yet were not considered with the findings from the test results.

Principals supported the purpose of high-stakes testing as the improvement of student achievement and indicated that although the results were consistently used for that purpose high-stakes testing did not improve education for all students. A majority of principals indicated that high-stakes testing appropriately held them accountable for student achievement as measured by the tests, but did not evaluate their school leadership abilities.

INDEX WORDS: High-stakes testing, Accountability, Student achievement.

AN ANALYSIS OF GEORGIA ELEMENTARY PRINCIPALS' PERCEPTIONS OF
HIGH-STAKES TESTING

by

SABRINA VAUGHN CALHOUN

B.S.ED., BERRY COLLEGE, 1978

M.ED., GEORGIA SOUTHERN UNIVERSITY, 1981

ED.S., GEORGIA SOUTHERN UNIVERISTY, 1991

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

DOCTOR OF EDUCATION

STATESBORO, GEORGIA

2007

© 2007

Sabrina Vaughn Calhoun

All Rights Reserved

AN ANALYSIS OF GEORGIA ELEMENTARY PRINCIPALS' PERCEPTIONS OF
HIGH-STAKES TESTING

by

SABRINA VAUGHN CALHOUN

Major Professor: Michael D. Richardson

Committee: T. C. Chan
Ababayehu Tekleselassie
Linda M. Arthur

Electronic Version Approved:
May 2007

DEDICATION

I would like to dedicate my dissertation to my Lord and Savior who has blessed me to complete this Doctorate, to my parents, Harold and Earline Vaughn, and my sons, Logan and Coleman. My parents reared a family of six children with their trust in God, one income, and my mother at home for us. They held high expectations for our behavior, consideration for others, and our efforts to always do our best.

ACKNOWLEDGEMENTS

I would like to acknowledge Dr. Michael D. Richardson for his expert guidance and continual encouragement during all of my work. He believed in me and taught me to believe in myself. I would also like to thank my other dissertation committee members, Dr. Ababayehu Tekleselassie and Dr. Linda M. Arthur, who gave their time and effort to assist in the development of this dissertation.

I would also like to acknowledge the love and support of my sisters and brothers, Cindy, Lisa, Tim, Mark, and Russ. I thank my nephew, Jordan Jones, for his daily encouragement. I thank my best friend since first grade, Joe Odom, and my church family at Cross Timbers Baptist Church for all of their prayers and encouragement.

TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS	7
LIST OF TABLES	11
CHAPTER	
1 INTRODUCTION	13
Introduction	13
Statement of the Problem	19
Research Questions	20
Significance of the Study	20
Limitations.....	22
Procedures	22
Definitions of Terms	25
Summary.....	27
2 REVIEW OF RESEARCH AND RELATED LITERATURE.....	29
Introduction	29
Role of Legislation	29
Role of Testing	32
Principal's Role	42
Summary.....	43
3 METHODOLOGY	45
Introduction	45
Research Questions	45

	Procedures	46
	Summary.....	55
4	REPORT OF DATA AND DATA ANALYSIS	57
	Introduction	57
	Research Questions	58
	Findings	59
	Summary.....	87
5	SUMMARY, CONCLUSIONS, AND IMPLICATIONS	89
	Introduction	89
	Research Findings	91
	Discussion of the Research Findings.....	92
	Conclusions	94
	Implications	96
	Recommendations	97
	REFERENCES	100
	APPENDICES	107
	A GEORGIA ELEMENTARY PRINCIPALS' PERCEPTIONS OF HIGH- STAKES TESTING SURVEY	108
	B STUDY PARTICIPATION INFORMED CONSENT LETTER.....	111
	C IRB PROPOSAL FORMAT FOR RESEARCH INVOLVING HUMAN SUBJECTS.....	113
	D IRB REVIEW BOARD APPROVAL FORM	117
	E FACE VALIDITY REQUEST	119

F	PILOT STUDY INFORMED CONSENT LETTER.....	121
G	STUDY PARTICIPANT INFORMED CONSENT LETTER.....	123

LIST OF TABLES

	Page
Table 1: Item Analysis	50
Table 2: Respondents' Years of Experience as Principal	60
Table 3: Principals' 2005-06 AYP Status.....	61
Table 4: Principals' Age Reported in Categories	61
Table 5: Principals' Gender	62
Table 6: Principals' Educational Level.....	63
Table 7: School Community Served.....	64
Table 8: Principals Should be held Accountable for Student Achievement as Measured by High-Stakes Testing.....	65
Table 9: Student Achievement will Improve as Schools Receive Rewards Based on High-Stakes Testing.....	66
Table 10: Student Achievement will Improve as Schools Receive Sanctions Based on High-Stakes Testing.....	67
Table 11: High-Stakes Testing is Fair	68
Table 12: High-Stakes Testing Improves Education for all Students.....	69
Table 13: High-Stakes Testing Results Appropriately Measures Accountability of Principal.....	70
Table 14: Principals Consistently Use High-Stakes Testing Results to Improve Student Achievement	71
Table 15: Factors beyond Principal Control Influence High-Stakes Testing and Should be Considered	72
Table 16: Purpose of High-Stakes Testing is to Make Principals More Focused Instructional Leaders	73
Table 17: High-Stakes Testing Results Accurately Evaluate Principals' School Leadership.....	74

Table 18: High-Stakes Testing Results Determine Principals' Perceptions of High-Stakes Testing	75
Table 19: Purpose of High-Stakes Testing is to Improve Student Achievement	76
Table 20: Purpose of High-Stakes Testing is to Hold Principals' Accountable for Performance	77
Table 21: Open-Ended Question 1 (Survey Question 14) How will high-stakes testing influence the way you hold yourself accountable for student achievement in your school?.....	78
Table 22: Open-Ended Question 2 (Survey Question 15) How has high-stakes testing impacted student achievement in your school?	82
Table 23: ANOVA Summary Table for Length of Service and Survey Questions.....	84
Table 24: ANOVA Summary Table for Age and Survey Questions.....	85
Table 25: ANOVA Summary Table for Gender and Survey Questions.....	85
Table 26: ANOVA Summary Table for Educational Level and Survey Questions	86
Table 27: ANOVA Summary Table for School Community Served and Survey Questions	86
Table 28: ANOVA Summary Table for AYP Status for 2005-06 School Year and Survey Questions	87

CHAPTER 1

INTRODUCTION

An elementary school principal commented on high-stakes testing:

A teacher knows that his whole professional status depends on the results he produces and he really is turned into a machine for producing those results; that is, I think, unaccompanied by any substantial gain to the whole cause of education. (Sutherland, 1973, p. 68).

Introduction

While it would be easy to assume that the above comment came from a modern day administrator, the quotation is actually taken from an account of a school master's writings in 1887. The reference is to the "payment by results" system for the public elementary schools in England and Wales. Beginning in 1862 and continuing for 30 years, the grants to each school were based on annual inspections of individual students. The students were tested on reading passages, arithmetic test cards, and other subjects later in this period (Jones, 2001). More than one hundred years later, high-stakes testing remains a debated issue in school reform, both in England and Wales, and now in the United States.

President George W. Bush signed the No Child Left Behind Act of 2001 (NCLB), on January 8, 2002, with the guarantee that no student would be left behind. There was strong public support for equality, periodic testing, highly qualified teachers, and the other provisions of the law. The promised benefits included that all students and subgroups of students would reach meaningful high standards, at the required Adequate Yearly Progress, AYP, pace. Principals, without consideration of their resources or the social capital of their students, are left with the daunting responsibility, to reach the high levels of student achievement as mandated by NCLB. The principal is the school-based

leader of the educational system which provides the framework for a community. The quality of life in that community depends on the development of the intellectual, emotional and social potential of each individual student. High-stakes testing is a key component of the NCLB, a law that stems from the basic assumption that every child - regardless of income, gender, race, ethnicity, or disability - can learn and deserves to learn. All efforts toward reforming schools must focus on ensuring the improvement of student achievement and learning. The NCLB Act represents the most sweeping change to the Elementary & Secondary Education Act (ESEA) since ESEA was enacted in 1965 (Peterson & West 2003). The current school reform mandate is for the principal to lead the school in the provision of a quality education for every student by the 2013-2014 school year. Accountability for this performance goal is based on student performance on high-stakes testing.

According to David S. Broder (2001), political analyst, NCLB “may be the most important piece of federal legislation in thirty-five years” (p. 31). The NCLB Act stands alongside the 1965 compensatory education law and the 1974 special education law in regard to the direction provided for federal, state, and local school spending. Every state, in order to receive federal aid under the terms of NCLB, must enact standards, together with a detailed testing plan designed to determine accomplishment of the standards. If a school fails to meet the standards, the students may transfer to another school in the same district. If a school consistently fails to make adequate progress, it becomes subject to corrective action (Peterson & West, 2003).

The goal of the NCLB Act is to provide a quality education for all students by the 2013-2014 school year (Georgia Department of Education [GDOE], 2003). This goal

builds upon a foundation of accountability for improved student achievement, increased flexibility and local control, expanded parental options, and data-driven research-informed instruction (Sclafani, 2002). Each state must define (AYP), a set of performance goals that establishes the minimum levels of improvement that must be achieved within specified time frames. And, like the payment by results of the late 1800s in England, the performance goals are based on student performance on state standardized tests (Mathis, 2003).

Today, much the same as in the 1800s, the utilization of high-stakes testing is a controversial issue in the theme of accountability (Hombo, 2003). As the goal of NCLB is a quality education for all students by 2014, the current measure of accountability for the goal is high-stakes testing. However, while educators acknowledge the concept of accountability, they are concerned about the demands of coercive accountability. Wayne Johnson, president of the California Teachers Association, indicated that the union was supportive of NCLB legislation. However, he indicated the union's concerns when he expressed concern about the use of standardized tests as the sole criteria for the determination of the accomplishment of public school students and teachers (Peterson & West, 2003).

According to Sclafani (2002), NCLB is based on four principles to guide school reform: (1) placement of responsibility on adults for the quality of learning in schools; (2) placement of control of schools in state hands; (3) respect for parental choice; and (4) utilization of scientifically-based research. Assessment is a key component of the plan to ensure that all students receive an adequate education (Sclafani).

Assessment is critical to making schools accountable and to identifying practices that make schools and teachers successful. The annual assessments mandated by NCLB ensure a rich data source, to be used to help individual children and to identify the strengths and weaknesses of students and teachers. Assessment of student learning is said to be a valuable tool for determining the professional development needs of principals and teachers. One interpretation of data gathered longitudinally from the test performance of students in a particular teacher's classroom could be that the teaching strategy or the content knowledge of the teacher could be a problem if the students are missing certain objectives every year (Sclafani, 2002).

The testing movement has become the central focus of education as an outgrowth of the present political focus on the inadequacies of public education and the lack of accountability for results (Kohn, 2000). Accountability has become one current avenue for politicians to express their concern about school achievement by getting tough with administrators, teachers and students. Despite the fact that test scores provide a quick-and-easy, and often, inaccurate way to chart school progress, politicians have added high test scores to the current political slogans like "tougher standards", "accountability" and "raising the bar" (Kohn).

The political focus has created legislative-centered classrooms rather than teacher-centered classrooms; testing classrooms rather than learning classrooms. The emphasis on high-stakes testing and tactics of control to manipulate the educational system to produce the higher scores has significantly impacted administrators, teachers and students (Kohn, 2000). As a rule, human beings are less likely to think creatively

when they perceive themselves to be under threat, and that is what many see as the current situation with its emphasis on higher test scores (Kohn).

Kohn (1993) indicated that rewards and punishments, one focus of NCLB, never produce more than temporary compliance, and even that is achieved at a substantial cost. Both represent ways of doing things to people, rather than working with people. It may be ineffective at best, to use rewards or punishments to motivate people to accomplish important goals such as school improvement. School improvement is a cause that is too valuable and worthy to be based on single criterion high-stakes testing results (Kohn, 2000).

When high stakes such as teacher salaries, school funding, student promotion/retention, college acceptance and job acquisition are attached to test results, the focus of education is on the high-stakes testing (Brown, 1993). Focus on a test will necessarily limit the information being taught. Any test can only contain a finite amount of information or number of items, regardless of the format. The motivation is to spend time and effort teaching the content on the test when the stakes are high for teachers, students and schools. According to Tunks (2001), the instruction that results from teaching to the test is misdirected and results in skills taught in isolation and a drill and practice type of instruction. The focus on curriculum is reduced to the parts targeted for testing, and other meaningful parts are disregarded.

There is also a higher likelihood that the integrity of the test will be compromised (Denig & Quinn, 2001). The pressure to produce high test scores may compromise administrations of high-stakes testing. Schools are ethical organizations and principals face ethical dilemmas on a daily basis. School leaders know that the decisions made, as

well as the values underlying those decisions, carry moral implications for the entire school community. Principals must be proactive and develop and defend principles of honesty and fair-testing practices for every student (Denig & Quinn). As principals address ethics and the accountability demands of high-stakes testing, they must consider the fairness of an approach that holds principals and teachers accountable for factors over which they have little control (Kohn, 2000).

According to Clancy (2000), the main information standardized test results provide is the amount of poverty and other associated variables in the communities where schools are located. The results may have very little to do with the measure of what happens in the classroom. The test results of a school primarily tell about the socioeconomic status and available resources in the homes of the community.

The leadership role of the school principal dictates early identification of struggling students and immediate action to help students achieve grade level skills based on the results of high-stakes testing (Sclafani, 2002). The principal must lead the school to produce student performance results on standardized tests to document the minimum levels of improvement that the school must achieve. The principal is held accountable for the adequate yearly progress of the school and that progress is based on the results of high-stakes testing. The school leader has to support, monitor, and assist teachers in teaching a rich curriculum to every child (Sclafani).

Statement of the Problem

The No Child Left Behind Act of 2001 (NCLB) is a comprehensive school improvement policy that mandates sweeping changes in longstanding educational practices. It puts the full weight of federal policy behind the accountability movement, mandating that schools bring all children, including minorities, English language learners, and the disabled to an adequate level of achievement.

Since one of the key components of the act is accountability, the assessment of student achievement has become critical as proof that all students are receiving an adequate education. As a determinant of school improvement, high-stakes testing is the current standard for accountability. Standardized testing is presently an annual event in Georgia and a critical determinant of accountability, funding, and the eventual survival or demise of a school.

Although principals are held responsible by NCLB for leadership decisions based on the results of the standardized tests, little is known about their perceptions of high-stakes tests and testing. Perceptions about the intent of the demands for the tests, the validity of the tests as a measure of school effectiveness, the purpose of the high-stakes testing and the expectations for the school leader may vary from principal to principal. The strategies principals utilize to interpret and communicate the results to improve student achievement are unknown.

Since there is little research to explain how principals perceive the demands of NCLB for high-stakes testing, the purpose of this study is to determine the perceptions of Georgia elementary principals about high-stakes testing. The study will address Georgia elementary principals' perceptions of the demands for high-stakes testing, the

appropriateness of the tests as a measure of school effectiveness, and basic premises of rewards and punishment to improve education.

Research Questions

The study is intended to answer the major research question: What are the perceptions of Georgia elementary principals about high-stakes testing? The following sub-questions will guide the research:

1. What do Georgia elementary principals perceive the purpose of high-stakes testing to be?
2. Do Georgia elementary principals perceive that high-stakes testing improves academic achievement?
3. Do Georgia elementary principals perceive that high-stakes testing appropriately measures the effectiveness of their performance?
4. What are the perceptions of Georgia elementary principals regarding rewards and sanctions for high-stakes testing?
5. Do demographic variables make any difference in principal's perceptions of high-stakes testing?

Significance of the Study

The principal plays a vital role in providing a quality education for all students by the 2013-2014 school year, the time line mandated by NCLB. It is important that principals are aware of the components of NCLB, the law that operates on the basic assumption that every child - regardless of income, gender, race, ethnicity, or disability - can learn and deserves to learn. The information gained from this study can provide

direction and strategies for those who are responsible for making decisions related to the provision of a quality education for every student.

There is a limited scope in the literature regarding the perceptions of elementary principals about the demands of high-stakes testing. Although studies about high-stakes testing are available, very few studies relate high-stakes testing to the accountability requirements of NCLB, or to the perceptions of elementary school principals. The results of the study will provide insight into the perceptions of elementary principals as they address the high-stakes testing demands related to the present mandates for accountability.

This importance of the research to State Boards of Education, State Departments of Education, and local Boards of Education is the use of high-stakes testing as the exclusive determinant of the accountability of elementary schools in Georgia and the nation. It is important that individuals in these positions understand accepted knowledge about learning capabilities, meaningful test scores, and school improvement.

This researcher is concerned with the current education reform mandate that all students reach meaningful high standards, at a required pace, determined by the score of one test that is administered annually. Education reform could benefit from an understanding of the perceptions of elementary principals about high-stakes testing. Principals make decisions based on their perceptions of these assessment results and the significance of the results. Assessment is a key component of NCLB which mandates that all students receive an adequate education. The assessment of students is critical for determining accountability of the school. Assessment serves the purpose of making schools accountable for student achievement.

The researcher will focus this study on the perceptions of Georgia elementary principals concerning high-stakes testing. It is important for a study of this type to be conducted now that Georgia educators, as well as national educators, are being held accountable to the standards set forth by NCLB. The high-stakes testing movement is not new to elementary principals in Georgia; however the impact of failing to show improvement is new. Principals may be removed from schools failing to demonstrate progress with the results of the high-stakes testing. Due to the lack of research on the perceptions of Georgia elementary principals concerning high-stakes testing, this study is needed to provide insight into the perceptions of others, and hopefully, to help them make the decisions required of them.

Limitations

Generalizability of study results will be limited specifically to Georgia elementary school principals. Applicability of the high-stakes testing components of the reform differ among elementary, middle and high school principals. The findings may be applicable to elementary principals in other states participating in similar mandates.

Procedures

Research Design

A quantitative approach was pursued in this study because it was an appropriate method for gathering descriptive information about the perceptions of elementary principals in Georgia about high-stakes testing (Creswell, 2003). The researcher's purpose was to survey a number of elementary principals and to generalize from a sample so that inferences can be made about the perceptions of the population. The survey was the preferred type of data collection procedure because of the economy of the design and

the rapid turnaround in data collection. The survey method was the preferred method in answering the research questions (Miller & Salkind, 2002).

Population

The researcher randomly sampled with replacement the 1,267 elementary school principals in Georgia. The principals were identified from the 2006 Georgia Public Education Directory published by the Georgia Department of Education.

Sample

A table of random numbers was used to select subjects from the elementary principals in the state of Georgia (Miller & Salkind, 2002). Currently 180 school systems exist in Georgia, and 1,267 principals were listed as heads of elementary schools containing grade groupings incorporating PK - 5 in various combinations (Georgia Department of Education, 2006). A total of 610 elementary school principals were selected for participation in the study. The sampling design for this population was a single-state sampling procedure because the examiner had access to the names in the population and could sample the people directly (Creswell, 2003).

Instrumentation

A survey was conducted to establish the statues of elementary principals' perceptions at a given point in time (Creswell, 2003). A self-designed survey questionnaire developed by the researcher explored principals' perceptions of high-stakes testing (Creswell). There have not been any surveys conducted in Georgia exploring principals' perceptions of high-stakes testing that were supported in the literature, and addressed unique facets of NCLB (Georgia Department of Education, 2003).

The survey instrument used to collect the data in the research study was developed by the researcher. An expert panel of judges was selected to examine the face validity of the instrument (Miller & Salkind, 2002). A cover letter (Dillman, 1978), the items and the closing instructions were included.

Pilot testing was conducted to establish the content validity of the instrument and to improve the questions, format, and the scale. Ten elementary principals from Alabama tested the instrument and their comments were utilized in the final instrument revision (Dillman). Data collected from the pilot study were also tested for internal consistency by using the Cronbach Alpha Test.

Questions for the survey were derived from a compelling interest on the part of the researcher to explore the perceptions of Georgia elementary principals about high-stakes testing. The survey included both open and closed form items to provide as through an investigation as possible, to understand the principals' perceptions as they participated in the implementation of the high-stakes testing components of NCLB.

Data Collection

The researcher collected data through the self-reporting of elementary principals' perceptions of high-stakes testing. The survey instrument used to collect the data in the research study was developed by the researcher. The survey instrument contained six questions addressing demographics, 13 questions about principals' perceptions of high-stakes testing and two open-ended questions. The survey was cross-sectional and the data was collected at one point in time (Miller & Salkind, 2002).

Data Analysis

The survey provided descriptive data that could be quantified and statistically analyzed with the SPSS 12.0 computer package (Statistical Package for the Social Sciences [SPSS], 2002). Descriptive statistics were employed to help summarize the overall trends and tendencies in the data, provide an understanding of the variability of the scores (Creswell, 2005). The researcher reported information about the number of members of the sample who did and did not return the survey. A table with numbers and percentages describing the respondents was used to present this information. The researcher provided a descriptive analysis of the data that indicated the means, standard deviations, and range of scores for these variables (Miller & Salkind, 2002). The two open-ended questions were used to obtain a narrative explanation of perceptions regarding the relationship of accountability and student achievement to high-stakes testing.

Definition of Terms

The terms used in this study are defined as follows:

Adequate Yearly Progress: A set of performance goals that establishes the minimum levels of improvement, based on student performance on state standardized tests, that schools, local education agencies, and the State as a whole must achieve within time frames specified in law.

Assessment: A primary tool for the reform of education which serves three fundamental purposes: the day-to-day management of instruction, the classification and placement of students, and the maintenance of accountability for educators and students.

Criterion-referenced test: A test that judges how well a test-taker does on an explicit objective relative to a predetermined performance level. There is no comparison to any other test-takers. The test tells how well students are performing on specific criteria, goals, or standards.

Educational Accountability: The process(es) by which school districts and states attempt to ensure that schools and school systems meet their goals.

Elementary and Secondary Education Act: Law enacted by Congress in 1965 to authorize and regulate the majority of federal K - 12 education programs to improve achievement among poor and disadvantaged students.

Elementary School: A school consisting of kindergarten through grade five or any combination of those grades.

High-Stakes Testing: The use of standardized tests to reward or sanction schools for their academic performance. It is among the most prominent education reform strategies. The idea behind high-stakes testing is that rewarding or sanctioning schools for their performance provides schools with incentives necessary to improve academic achievement.

Improving America's Schools Act of 1994 (P. L. 103-382): Reauthorized the Elementary and Secondary Education Act of 1965 (ESEA), with a focus on changing the way education was delivered, encouraging comprehensive systemic school reform, upgrading instructional and professional development to align with high standards, strengthening accountability, and promoting the coordination of resources to improve education for all children.

No Child Left Behind Act (NCLB): A law that operates on one basic assumption: that every child - regardless of income, gender, race, ethnicity, or disability - can learn, and that every child deserves to learn.

Perception: The representation of what is perceived; basic component in the formation of a concept, a way of conceiving something.

Principal: Any person holding the top administrative position in a school.

Standardized Testing: Testing with explicit, fixed procedures for administering, scoring, and interpreting the test. The test has been standardized or normed on a large, representative sample of individuals at specified grade levels. The test is standardized both in the sense of a common procedure for administering the test and common norms for interpreting the results.

Summary

High-stakes testing is presently a key component of NCLB, a law that stems from the basic assumption that every child - regardless of income, gender, race, ethnicity, or disability, can learn and deserves to learn. The goal of NCLB is to provide a quality education for all students by the 2013-2014 school year. The goals build upon a foundation of accountability for improved student achievement, increased flexibility and local control, expanded parental options, and data-driven research-informed instruction. Each state must define AYP, a set of performance goals that establishes the minimum levels of improvement that must be achieved within specified time frames. The performance goals are based on student performance on state standardized tests.

Educators acknowledge the concept of accountability but they are concerned about the demands of the current coercive accountability. The emphasis on high-stakes testing

scores has significantly impacted administrators, teachers and students. The principal must lead the school to produce student performance results on the standardized tests to document the minimum levels of improvement that the school must achieve in order to avoid restructuring.

There is little research to explain how principals perceive the demands of NCLB for leadership decisions, based on the results of standardized tests, to lead the school to the attainment of AYP each school year. Therefore, the researcher's purpose for this study is to determine the perceptions of Georgia elementary principals about high-stakes testing.

Principals are responsible for the accountability of high-stakes testing. Principals have tremendous responsibilities which relate to ethical issues surrounding high-stakes testing as well as the early identification of struggling students and the determination of immediate action to help the students achieve grade level skills.

CHAPTER 2

REVIEW OF RESEARCH AND RELATED LITERATURE

Introduction

Since its earliest beginnings, public education has been considered one of the finest accomplishments of the United States. Public schools were believed to provide opportunities for even the most naive students to achieve success, happiness, and a good life. This positive regard for public education, teachers, and the teaching profession continued well into the 1960s and was based on the assumption that students were learning, and teachers were teaching (Popham, 2001; Finn, 1991).

However, during the late 1960s and early 70s, public discontent for education first became evident as newspaper articles began to appear about students receiving high school diplomas yet being unable to properly complete job applications. Stories were published about students being socially promoted who had only rudimentary level reading and writing skills. The United States public schools and the nation's teachers began to face increasingly serious attacks (Popham, 2001).

Role of Legislation

As widespread citizen distress often engenders a legislative response, a number of state legislatures and state and district school boards soon established basic-skills testing programs commonly referred to as minimum competency tests which focused on reading, writing, and math. Students were required to pass the tests before they could receive a high school diploma. Students in lower grades were required to pass the tests before they could be promoted to the next grade (Benveniste, 1985; Finn, 1991).

Policymakers claimed the objective of minimum competence testing was to guarantee parents those students who passed the competency tests had mastered the basic skills measured by the tests (Popham, 2001; Ohanian, 1999). Although the focus appeared to be on students, the legislators and educational policymakers were actually looking at teachers who it was assumed, must be falling down on the job of teaching since children were making their way through the educational system without having learned how to read, write, or fill out a job application (Popham, 2001; Sclafani, 2002).

The establishment of minimum competency testing programs was firmly supported by members of the business community. Corporate America needed high school graduates with basic reading, writing, and math skills. Hence, corporate America endorsed competency tests based on the premise that such tests could guarantee graduates possessing these skills (Popham, 2001; Schwartz & Gandal, 2000).

With the concentration on the shortcomings of education, considerable attention was focused on the disparity in educational opportunities and student performance. Hence, accountability systems were employed during the 1960s to hold schools, principals, teachers, and sometimes students responsible for education (Benveniste, 1985; Linn, 2000). Perhaps the most important educational legislation and accountability system passed at the time, The Elementary and Secondary Education Act of 1965 (ESEA), was signed into law on April 1, 1965 by President Lyndon Johnson.

ESEA was the first major and the most enduring federal legislation to improve education (U.S. Department of Education, 2001; Lemann, 2000; Linn, 2000). ESEA was designed to provide significant financial resources for educationally deprived children and to become the keystone to President Johnson's "War on Poverty" (Webb, et al,

1992). As a result of ESEA, (Yudof, Kirp, & Levin, 1992), educational funding moved from general to categorical and was tied to national policy. Funds were assigned based on the student's poverty status rather than the type of school, and the allocation of funds to the state department of education created an increase in the state bureaucracy's role in educational decision-making (WestEd, 2000). Educators were required to annually evaluate their efforts for the first time (Worthen, Sanders, & Fitzpatrick, 1997). The testing of students in grades three through eight was mandated.

The ESEA Act nurtured the notion that students' test scores could determine a school's quality. Significant amounts of money were dispensed to school districts to support locally designed programs designed to improve student achievement. The amount of federal money allotted to school districts had been relatively modest before the 1965 enactment of ESEA (Popham, 2001). With the increased funding, corresponding safeguards were built in to monitor the new federal funds for educational initiatives. Robert Kennedy, then a senator from New York, championed the addition to the law that required educators receiving ESEA money to prove that the funds were well spent by evaluating and reporting on the federally funded programs (Popham).

The establishment of ESEA supported the public's perception that the quality of a school could be determined by its students' test scores. Educators receiving ESEA awards searched for suitable tests to measure student achievement as documentation of the success of their ESEA-funded programs. The Metropolitan Achievement Tests and the Comprehensive Tests of Basic Skills were the most readily standardized achievement tests at this time. These tests were frequently selected by school districts because they were developed by respected testing companies, subsequently regarded as technically

sound, and readily available. However, at the present time, there are still many questions about the depth of curriculum standards and the criterion-referenced tests which are reportedly aligned to the curriculum (English, 2003; Finn & Walberg, 1994). Educators knew little about test development and their work depended on ESEA dollars, so most educators readily accepted the standardized achievement test results as accurate depictions of the quality of classroom instruction. However, the measures were not strongly related to the skills and knowledge being taught by any particular classroom program funded by ESEA (English & Hill, 1994). Even so, the use of standardized achievement test scores as determinants of the success of ESEA programs became the national norm (Popham, 2001).

Role of Testing

The seemingly next logical assumption was made by policymakers and educators. If standardized achievement tests could determine the effectiveness of the basic skills-focused instructional programs funded by ESEA, they could also be used to evaluate the effectiveness of other instructional programs. Although the 1965 ESEA stimulated the increased reliance on standardized achievement tests to judge educational success, educators allowed their teachers to be evaluated by students' scores on the tests (Popham, 2001).

The minimum competency tests of the 1970s and early 80s focused on remarkably low level skills and knowledge as a result of the establishment of statewide minimum competency testing programs as mandated by state laws passed at the time. The law's implementation was usually turned over to the state's education department. The officials of the education department created the authorized assessment programs which

were constructed by external test development firms as few state departments of education possessed the ability to generate high-stakes assessments (Popham, 2001).

State officials then determined the skills and knowledge to be measured. These decisions were made by a committee of state educators appointed by state authorities. The committees operated under the guidance of the external staff and identified the skills and knowledge in each subject content area that would be measured by the competency test. The committees often selected very low-level basic skills and knowledge. Most states ended up with minimum competency tests because committees chose low-level content to be assessed. They realized that denying diplomas and grade promotions to students because of low test scores would reflect negatively on teachers.

Once the minimum competency test was in place, teachers devoted considerable energy to teach students to master its content and pass the tests. Relatively few students failed the early minimum competency tests and those who did were provided opportunities to re-take the tests. At least some students did fail and were denied diplomas (Popham, 2001).

Newspaper writers quickly discovered that something newsworthy was happening. They could easily write stories that compared schools within a district on the basis of test failure rates and subsequent diploma denial. With the reading of the articles, a public perception began to emerge that good school had few students to fail and bad schools had many students to fail. The quality of a school was linked to the quality of the test scores of the students. This approach to judging schools flourished (Popham, 2001). The standardized test results often exercised a powerful negative influence on teachers and a

few points gained can have a tremendous impact on the public image of schools and teachers (Finn & Walberg, 1994).

In the 1983 written report, A Nation At Risk, it was claimed there was a significant decline in education in the United States (National Commission on Excellence in Education, 1983). The report was released by the Reagan White House on April 26, 1983. Citizens, bureaucrats, and business leaders were alarmed by the findings that American schools were in crisis and public education was in decline (National Commission of Excellence in Education). Educators were criticized and scrutinized. New reform initiatives were discussed and included accountability through performance rating, incentive systems, and changing professional roles (Utah Foundation, 1999). The adoption of National Education Goals led to a greater acceptance of this approach to education reform as the goals required content and standards (Finn & Walberg, 1994). However, the same national education goals were highly criticized because the means to accomplish them were not identified (English & Hill, 1994).

President George H. Bush and fifty governors declared that every American child should meet challenging academic standards by the year 2000 at the Charlottesville Education Summit in 1989 (Finn, et al., 1998; Hadderman, 2000). President Bush supported the establishment of national standards for all American children to assure that students were receiving a quality education. A key development was the adaptation of President Bush's and the nation's governors' six national education goals in 1989. Efforts were made to establish national standards and tests (Hadderman, 2000).

Standards are “clear articulation of what students should know and be able to do” (Reeves, 2000, p. 5). Ohanian (1999) indicates that a standards based curriculum based on the assumption that one curriculum is appropriate for everyone devalues children. According to Finn (2000), the setting of standards – student expectations – is the first step towards school reform. High-achieving successful school systems in other industrialized nations have national education standards and performance standards that determine mastery (Finn & Walberg, 1994). The initiative for higher standards and stronger assessment instruments is viewed as an appropriate strategy for improving student academic achievement (Schwartz & Gandal, 2000). However, the attempts to write national standards in the United States failed. Despite this failure, states were encouraged to develop demanding content and performance standards as a central part of state reform efforts (Linn, 2000). At the end of the 1990s, many states had but many states were still in the process of writing state standards and developing functional accountability systems for their school (Finn, et al., 1998).

Reform initiatives became categorized as “the era of legislated learning” (Glickman, 1990, p. 38) as the education reform movement moved into the 1990s. The minimum-competency testing movement of the 1970s was intensified in the late 1980s and early 1990s. The use of standardized test results for accountability purposes was expanded (Linn, 2000). According to Linn (2000), “accountability programs took a variety of forms, but shared the common characteristic that they increased real or perceived stakes of results for teachers and educational administrators” (p. 7).

School reform legislation and educational accountability continue to be primary issues of federal and state governments. Under the leadership of President George W.

Bush, the federal government established its assurance of accountability for student achievement by enacting NCLB, which became law on January 8, 2002. The NCLB Act reauthorized ESEA to provide all American school children with the opportunity and means to achieve academic success (U.S. Department of Education, 2001).

A bipartisan effort, NCLB represents the greatest federal incursion into K-12 education to date (Gallup & Rose, 2002). The plan extended the federal government's decision-making role regarding K-12 schools and called for increased federal attention to standards, assessments and accountability, and implementation of federal government priorities (Education Commission of the States, 2001). Federal rewards and sanctions based on meeting acceptable levels of educational change as outlined in the plan were specified, while the power to continue to develop independent accountability systems was left to the states (Education Commission of the States).

When President George W. Bush signed NCLB into law, he made the national government a prominent player in the effort to use high-stakes accountability to drive school improvement (Peterson & West, 2003). Accountability is the underlying theme of educational reform and educational assessment is the tool that serves to evaluate and make judgments regarding student knowledge and understanding (Bernhardt, 2001). There are many forms of assessments, but tests have emerged as the most preferred way of measuring student achievement in accountability systems (Bernhardt). Even as studies predict that a focus on test results alone does not improve education, testing continues to receive the most attention and has had the most impact of all the elements in accountability systems (Bernhardt; English & Hill, 1994).

Educational testing has been defined in numerous ways. A standardized test is any test that is administered and scored in a standard, predetermined manner (Popham, 2001). According to Kerlinger (1986), standardized tests are published group tests that are based on general educational content common to a large number of educational systems. They are the products of a high degree of professional competence and skill in test-writing and are usually quite reliable and valid. A test as a systematic procedure in which the individuals tested are presented with a set of constructed stimuli to which they respond. The responses enable the tester to assign the student numerals or sets of numerals from which inferences can be made about the students' possession of skills and competencies the test is supposed to measure. In other words, a test is a measurement instrument (Kerlinger).

According to Lyman (1998), a standardized test is an empirically developed test, designed for administration and scoring according to stated directions, for which there is evidence of validity and reliability, as well as norms. Achievement tests measure present proficiency, mastery, and understanding of general and specific areas of knowledge. For the most part, they are measures of the effectiveness of instruction and learning and enormously important in education.

Testing and assessment results are the greatest single source of knowledge about academic achievement (Finn, 1991). They can be used to provide performance data about the student, classroom, school, local system, state and nation as a whole. However, pains should be taken to assure that the data derived from testing programs should be timely and trustworthy. Testing is an immense topic with critical elements. It is crucial to obtain assessment instruments that measure that which is considered important for

children to learn (Finn). Tests should be carefully aligned to curriculum objectives, instructional materials, and pedagogy but this seldom occurs in American education.

According to Heubert and Hauser (1999), tests must meet professional standards of reliability, validity and fairness especially when they are used for making high-stakes decisions such as promotion and graduation for individual students (English, 2003; Fullan, 2005). Public support continues at high levels for making such decisions, even if the use of high-stakes testing does lead to lower rates of promotion and high school graduation (Johnson & Immerwahr, 1994; Hochschild & Scott, 1998).

The National Academy of Sciences, through its National Research Council, determined that the use of high-stakes testing for promotion, tracking or graduation is intended for the purpose of setting high student learning standards, and raising achievement levels, parental involvement and public support for education (Heubert & Hauser, 1999). But the intended benefits of high-stakes testing must be weighed against the unintended negative consequences for some students. Also, the costs and benefits of testing should be balanced with making such decisions using criteria other than test scores. Hence, a comprehensive high-stakes testing policy should include concern for balance between the benefits and costs to the individual as well as for the mass (Heubert & Hauser).

According to Peterson & West (2003), high-stakes accountability represents a compromise among policymakers for addressing what schools should focus on and how school performance should be measured. The advantages include specification of skills and knowledge to be mastered which provides educators with clear educational goals. According to Finn & Walberg (1994), high-stakes testing teaches students to handle

pressure, deal with failure and gain self-esteem from one's own academic achievement. These are important parts of growing up and things principals and teachers should help students learn in school. High-stakes accountability can improve the professionalism of educators and improve public support of education as educators are held to clear standards, and those educators who do not meet the standards can be sanctioned (Peterson & West).

According to Finn (1991), accountability systems are constructed to get good results and there is a greater likelihood of the production of good results if they are followed by consequences. American education has evolved into an outcomes based project where everyone will be held responsible by the public for their results. This is the only accountability system that can adequately address the weak academic achievement of students. Until now, schools and educators have been judged for a very long time on an accountability system based on their compliance with rules, procedures and allocations, which were believed to be the proper criteria for judging their work (Finn).

The testing component of the current accountability system is an immense topic, and there are appropriate criticisms of tests (Finn, 1991). Deming indicates that to begin to address school reform, the idea of inspecting students must first be confronted as the hallmark of the accountability system calls for more and more testing (English & Hill, 1994). Ohanian (1999) questions the need for more tests when every teacher and student already knows which students are performing below grade level. Sometimes tests are used for purposes other than they were designed, and sometimes tests do have objectionable side effects. The crucial point is to utilize an assessment instrument that measures information determined to be the most important for students to learn. The

subject matter that is tested should be what schools emphasize and what students learn. Hence, it is appropriate to expect that the testing program used in schools be carefully aligned with the curriculum objectives and instructional materials (English, 2003; English & Hill, 1994). Ensuring that students learn the content of a test that does a good job of probing the knowledge and skills deemed important for children to acquire, is the proper function of testing and classroom instruction within an accountability system (Harris, 1989).

High-stakes tests based on clearly articulated standards represent student performance and hold students, teachers, and schools to the same challenging standard (WestEd, 2000). High-stakes tests highlight the achievement gap between rich and poor, school systems and ethnic and cultural groups according to. The data forces local educators to address this problem through the rewards and sanctions awarded based on overall student achievement. The assessment results can be used to determine if all students are mastering the knowledge and attaining the skills necessary for academic success. The results of high-stakes testing can help improve student achievement when used for appropriate purposes (American Educational Research Association, 2000).

A great deal of research has shown that students master knowledge and attain skills from what they are taught (Porter, 1998). Thus, an important determinant of the fairness of a high-stakes test such as a graduation test is the degree to which curriculum and instruction are aligned with that measured by the test. It is challenging and expensive to measure the content of actual instruction (Popham & Lindham, 1981). As a result, it appears that many states have not properly validated their assessment instruments, and that proper studies would reveal important weaknesses (Stake, 1998).

Lyman (1998) reports that there are still frequent and virulent attacks on testing, but people are using them more today than ever before. He cites criticisms including labeling of children, invasion of privacy, inconsistent results, unfairness, and gross misinterpretations.

Some of the most frequent criticism of testing includes the excessive reliance on multiple-choice formats and standardized tests, the irresistible temptation to cheat on high-stakes tests, and biased and discriminatory results (Finn, 1991). Other criticisms include the fact that some children do not do well on tests, and it is unfair to make such important decisions such as promotion and high school graduation on the basis of a test score. Tests indicate cumulative achievement rather than how much a student has learned in a specific classroom or school, and they take up too much time that should be devoted to teaching (Finn).

High-stakes tests narrow the curriculum and instruction due to the difficulty in addressing important curriculum goals that require generative thinking, sustained effort over time, and effective collaboration. High-stakes testing may lead to retentions and the withholding of high school diplomas even though all curriculum goals are not included in assessments. The immediate effect on students and the impact on society over the long term must be considered (WestEd, 2000). Additional testing can seriously burden teachers and students and further limit already precious instructional time and resources.

Additional criticism on high-stakes testing includes, according to Bracey (2000), the instruments and technology have not met the demands placed upon them for high-stakes accountability in most cases. The assessment systems designed to monitor student achievement lose dependability and credibility when high stakes are attached to the

systems. Often, the unintended negative effects of high-stakes testing outweigh the intended positive effects. As high-stakes testing becomes more prevalent, the debate continues as to whether the testing supports the reform of instruction or leads to limited instruction and student achievement (Firestone, 2004).

With regard to instruction and student achievement, state testing programs are causing teachers to change what they teach and how they teach it, according to a report by Pedulla, Abrams, Madaus, Russell, Ramos, and Miao (2003). According to the results of the survey, teachers in high-stakes situations reported feeling greater pressure to have their students to perform well, instruction to be aligned with the test, and engage more in test preparation. Approximately seventy-five percent of teachers reported intense pressure to perform well and extreme anxiousness about taking the state test. Teachers found that the benefits of testing were not worth the time and money involved. Elementary teachers indicated that their teaching contradicted their ideas of sound instructional practices.

Principal's Role

The school principal is pivotal to the success of educational accountability in schools (Fullan, 2003). The success of school improvement and accountability relies heavily on the leadership skills of high quality principals for reaching instruction and learning goals (Leithwood, Seashore, Anderson, & Wahlstrom, 2004); Waters, Marzano, & McNulty, 2003). The basic work of quality principal leadership is to ensure improved teaching and increased student achievement. Michael Fullan wrote:

Leaders have a responsibility to invest in the development of organizational members, to take the chance that they will learn, and to

create environments where people will take risks, tackle difficult problems, and be supported in this endeavor . . . In many ways, accountability is built into the culture. The day-to-day interaction among peers and between peers and others creates a system of checks and balances of learning and accountable performance (p. 43).

Improvement in student achievement and overall school improvement is mandated to be accomplished in public schools throughout the state of Georgia in NCLB. Principals, as school leaders, have the responsibility for achieving this improvement.

Summary

Research revealed that public discontent for education first became evident during the late 1960s and early 1970s as newspaper articles began to appear about students receiving high school diplomas without being able to properly complete job application forms. The legislature responded to the widespread citizen distress with legislative responses establishing basic-skills testing programs in reading, writing, and math for promotion in the lower grades and for a high school diploma.

Considerable attention was focused on the disparity in educational opportunities and student performance with the concentration of the shortcomings of education. Accountability systems were employed to hold schools, principals, teachers, and sometimes students responsible for education. The Elementary and Secondary Education Act of 1965 (ESEA) was the first major and the most enduring federal legislation to improve education. Funds were assigned based on the poverty status of students and educators were required to annually evaluate their efforts for the first time.

The seemingly next logical assumption was made that if standardized achievement tests could determine the effectiveness of basic-skills programs for funding purposes, they could also be used to hold principals and schools accountable for improving student achievement. The positive aspects of high-stakes testing include the provision of performance data about the student, classroom, school, locally system, state and nation as a whole. However it is crucial to obtain assessment instruments that measure that which is considered important for children to learn and the high-stakes tests should be carefully aligned to curriculum objectives, instructional materials, and pedagogy. The intended benefits of high-stakes testing must be weighed against the unintended negative consequences for students. State testing programs are having tremendous impact on what teachers teach and how they teach it. The school principal is pivotal to the success of educational accountability in school.

CHAPTER 3

METHODOLOGY

Introduction

The No Child Left Behind Act (NCLB) of 2001 reauthorized The Elementary and Secondary Education Act (ESEA) and has significantly raised expectations for states, local school districts, and schools in that all students will meet or exceed state standards in reading and mathematics within twelve years. Principals, as leaders of schools, have been delegated the responsibility to influence their stakeholders to meet the demands of the state academic standards and testing system to meet federal requirements. However, comprehensive education reform efforts may result in less than the desired outcomes (Fullan, 1993). Problems including confusion, ambiguity, and conflict can occur when new structures are quickly created to implement the latest policy. Meaningful education reform depends on the ability to survive through the planned and unplanned changes that occur while a school grows and develops (Fullan).

The facilitation of school reform efforts such as the Adequate Yearly Progress (AYP) accountability mandates have influenced the ways in which principals structure their schools, deliver the curriculum, and develop their collaborative relationships with stakeholders (Cousins & Leithwood, 1992). Georgia elementary principals' perceptions of their duties and responsibilities have provided worthy data for documentation of perceptions of high-stakes testing during the implementation of NCLB.

Research Questions

The major research question was: What are the perceptions of Georgia elementary principals about high-stakes test? The following sub-questions guided the research:

1. What do Georgia elementary principals perceive the purpose of high-stakes testing to be?
2. Do Georgia elementary principals perceive that high-stakes testing improves academic performance?
3. Do Georgia elementary principals perceive that high-stakes testing appropriately measures the effectiveness of their performance?
4. What are the perceptions of Georgia elementary principals regarding rewards and sanctions for high-stakes testing?
5. Do demographic variables make any difference in principal's perceptions of high-stakes testing?

Procedures

Research Design

A descriptive, survey approach was employed to address the research questions. The researcher's first purpose for the utilization of the survey design was for consideration of problems or current conditions and practices (Creswell, 2003) relating to elementary principals' perceptions regarding effects of high-stakes testing on their roles. The researcher's second purpose was to examine differences among groups of principals experiencing similar difficulties with accountability especially as it related to implementation of NCLB high-stakes testing mandates and to identify ways that their experiences could be informative in future decision making and planning (Creswell).

Population

The population for this research was comprised of 1276 Georgia elementary school principals currently listed in the Georgia Public Schools Directory (Georgia Department

of Education, 2006). Special entities such as specialty schools, alternative schools, psychoeducational programs, etc. were listed in the directory, but were excluded from the study. Some school systems divided the elementary grades into primary and intermediate levels. These schools were included according to the written designation provided by the system, and listed as such in the Georgia Public School Directory (Georgia Department of Education, 2006).

Sample

A table of random numbers was used to select subjects from the elementary principal population in the state of Georgia (Gall, Borg, & Gall, 1996). Currently 180 school systems have been found to exist in Georgia, and 1276 principals were noted as heads of elementary schools containing grade groupings incorporating PK-5 in various combinations (Georgia Public School Directory, 2006). A random sample with replacement of elementary principals was selected to ensure that each Georgia elementary principal would “have an equal and independent chance of being selected as a member of the sample” (Gall, et al., 1996, p.223). This method ensured that a sufficient representative sample of principals was used for the purpose of generalizing the results of the study to the target population (Creswell, 2003). A single-stage sampling procedure was used since principals’ names and school addresses were listed in the directory.

A total of 610 elementary school principals were selected for participation in the study. This number was slightly more than double the recommended sample size of 302 for a total population size of 1350 (Krejcie & Morgan, 1970). This has taken into consideration of approximately 50% return rate.

Instrumentation

A cross-sectional survey was conducted to establish the status of elementary principals' perceptions at a given point in time. A survey was developed by the researcher to collect data in the investigation of principals' perceptions of high-stakes testing (Babbie, 2001). There have not been any surveys conducted in Georgia examining principals' perceptions of high-stakes testing as related to NCLB (Georgia Department of Education, 2003). The researcher selected survey questions that reflected components of high-stakes testing that were supported in the literature, and addressed unique facets of NCLB (Georgia Department of Education).

The researcher designed questions that were derived from a compelling interest on the part of the researcher to explore the perceptions of Georgia elementary principals about high-stakes testing. The complete survey, Georgia Elementary Principals Survey on High-Stakes Testing, can be found in Appendix A. The researcher included quantitative and qualitative questions, for the purpose of deriving a thorough understanding of principals' implementation of the high-stakes testing components of NCLB. The researcher addressed what elementary principals perceived the purpose of high-stakes testing to be and if they perceived improved academic performance. In addition, the researcher endeavored to ascertain whether principals perceived that high-stakes testing was an appropriate measure of the effectiveness of their performance and if the use of rewards and punishments improved education. Further, the researcher endeavored to compare demographic variables that might affect survey responses.

The researcher extensively reviewed the research literature for development of specific survey questions. Unique facets of NCLB were included to answer the research

questions. The researcher reported the research questions and the subsequent relationship with research literature, and survey questions in Table 1.

Table 1

Item Analysis

Survey Question Number	Alignment with Literature	Research Question Answered
1	Denig, 2001; Sclafani, 2002; Hombo, 2003	3
2	Denig, 2001; GaDOE, 2003; Mathis, 2003	4
3	Denig, 2001; GaDoe, 2003; Mathis, 2003	4
4	Brown, 1993; Hombo, 2003; Mathis, 2003	Major
5	Hombo, 2003; Mathis, 2003;	2
6	GaDoe, 2003; Mathis, 2003	3
7	Denig, 2001; Sclafani, 2002	2
8	Denig, 2001	Major
9	Jones, 2001; Mathis, 2003	1
10	GaDoe, 2003; Hombo, 2003; Mathis, 2003	3
11	Clancey, 2000; Mathis, 2003	Major
12	Sclafani, 2002; Mathis, 2003	1
13	Clancey, 2000; Sclafani, 2002	1
14	Clancey, 2000; Jones, 2001	3
15	Tunks, 2001; Mathis, 2003	2

An expert panel of judges was empowered to examine the face and content validity as well as ease of completion of the instrument (Fowler, 2002). A committee composed of Dr. Tak C. Chan, Professor of Kennesaw State University; Dr. Jim Smith, College Dean of Bowling Green State University, Mr. Gene Baker, Principal of Rockdale County; Dr. Cindy Reed, Auburn Department Chairman, Principal, and Dr. Maggie Shook, Superintendent of Prospect, Connecticut Region 16 Board of Education, served as the expert panel.

After an affirmative response from each panel member, a letter was mailed on May 15, 2006 (see Appendix E), formally requesting his or her expertise in determining the face validity of the survey. The mailing included the cover letter clarifying the research questions driving the survey, the survey itself, and a self-addressed stamped envelope. Recommendations for survey improvement were requested and panel members were asked to make notations regarding the amount of time required to complete the survey. Responses were requested with a 2-week time frame. All responses were received by June 5, 2006. The survey was then revised per recommendations of the panel.

A pilot test was conducted to assess how well the survey addresses issues examined in the study (Gall, Borg & Gall, 1996). Ten Alabama public school elementary principals were selected to respond to the survey in order to determine the content validity of the instrument. The researcher requested that the pilot participants respond by writing their perception of what each question was asking beside the question. Ten responses were received. After the responses were received, the researcher was enabled to determine the accuracy of the wording of the survey questions and to ensure that the

intended purpose of each question would be conveyed to the survey respondents (Gall, et al.). The principals confirmed 15-20 minutes for completing the survey.

The questionnaire format included closed form and open-ended questions. It was constructed as an attitude scale with 15 items that required 15 responses, for the purpose of representative assessment of principals' perceptions (Gall, et al, 1996). The benefits of utilizing mailed attitude scales include greater coverage, which may yield greater validity through larger and more representative samples, a sense of privacy for respondents, and time for more considered answers (Wallace, 1954). There is greater uniformity of stimulus and thus greater reliability is achieved. The anonymity may encourage honesty and frankness. The mailed self-administered scale is economical (Kerlinger, 1986). A Likert scale and written responses were utilized to assess principals' perceptions (Gall, et al.).

The closed form questions were included for ease of response. There have been advantages attributed to the utilization of survey research rather than the interview method (Kerlinger, 1986). The interview is a special event in the life of the respondent and may affect the respondent so that he talks to, and interacts with, the interviewer in an unnatural way. The respondents may give answers that reveal what they would like to perceive rather than what they actually perceive. Interviewers must be trained, validity is a problem, and special pains must be taken to eliminate interviewer bias. Open-ended questions were included in the survey to capture the principals' own words, the perceived effects of high-stakes testing (Kerlinger).

Procedures

Permission to conduct the study was requested by the researcher from the Georgia Southern University Institutional Review Board. Receipt of permission to proceed was received on April 20, 2006. The preselected panel of experts was subsequently contacted by mail.

Each panel member was contacted in May, 2006 by the researcher, to request assistance with the study. It was explained that approval for the research proposal from the Georgia Southern University Institutional Review Board (IRB), had to be obtained prior to the beginning of the face validity phase of the study. The research proposal was submitted to the IRB (see Appendix C), and approval for conducting the study was received on April 20, 2006 (see Appendix D).

Initially, each member of the expert panel was contacted by the researcher requesting his or her participation in reviewing the survey. After receiving consent from the expert panel, the researcher sent a letter with a copy of the survey to each individual on May 5, 2006. The researcher requested a 2-week deadline. When the responses were received, the survey instrument was revised to reflect suggestions offered by the experts. The survey was revised for clarity based on the thoughtful reflection provided from each expert panel member.

A pilot test was conducted to assess how well the survey addresses issues examined in the study (Gall, Borg & Gall, 1996). Pilot testing was conducted to establish the content validity of the instrument and to improve the questions, format, and the scale. survey respondents. The selected pilot survey participants were contacted by mail with a cover letter (see Appendix F) explaining the study, specific instructions for evaluating the

survey, a copy of the survey, assurances that their identities would remain anonymous and that their specific responses would remain confidential on June 15, 2006. The researcher requested a 2-week deadline. After responses were received, no additional corrections were needed. The pilot study confirmed that each question was interpreted as intended by the researcher, and thus should be likewise interpreted by the selected study participants.

Data collected in the pilot study were also analyzed by using the Cronbach Alpha test to establish reliability. A Cronbach Alpha of .984 indicated that the survey instrument has a high degree of internal consistency.

The survey instrument with a cover letter and a self-addressed stamped envelope was mailed to the randomly selected elementary principals on November 21, 2006. The randomized number was printed on the front of each envelope. The randomized numbers corresponded to a list with the names of the selected principals. This number enabled the researcher to identify participants for the purpose of making contact with nonrespondents.

A second mailing of a revised cover letter survey and self-addressed envelope was required. A total of 200 surveys were mailed on December 28, 2006.

The surveys were received and tabulated by the preassigned number. The envelopes were destroyed as specified, to ensure confidentiality. A total of 335 surveys were used in the study. The targeted sample size of 302 was obtained and responses received represented 55% of the total number of contacts.

Data Analysis

The survey provided descriptive data that was quantified and statistically analyzed with the SPSS 12.0 computer package (Statistical Package for the Social Sciences [SPSS], 2003). The researcher provided a descriptive analysis of data in the study. The analysis indicated the frequencies, means, standard deviations, and range of scores for these variables (Miller & Salkind, 2002; Creswell, 2005). Selected NCLB law components and demographic variables included in the seventh research question (the relationship between demographic variables and principal response) were analyzed using the ANOVA components of SPSS (SPSS). The researcher used these methods of analysis to determine whether specific groups of elementary principals responded predictably to particular survey questions, and whether any differences found among groups were statistically significant (Gall, et al., 1996).

The two open-ended questions were used to obtain a narrative explanation of principals' perceptions regarding the perceived influence of high-stakes testing on the accountability of the principal for student achievement, and the impact of high-stakes testing on student achievement.

Summary

The researcher examined Georgia public school elementary principals' perceptions of high-stakes testing to determine their perceptions of the NCLB mandates. Personal and demographics that included years of experience, AYP status of 2005-06 school year, age, gender, degree, and school community setting were examined to compare groups of respondents.

After receiving permission to conduct the study, face validity of the survey was ascertained by a panel of expert judges. A pilot survey was conducted for the purpose of determining content validity of the survey, with 10 preselected Alabama public school elementary principals. Ten respondents participated in the pilot study phase of the research.

The data were analyzed utilizing quantitative and qualitative methods. Six demographic variables were analyzed utilizing the one-way analysis of variance statistical procedure. The software program SPSS (SPSS, 2003) was employed to analyze the data quantitatively. Open-ended questions were analyzed for patterns, themes and categories (Creswell, 2003).

Georgia elementary principals were provided a unique opportunity to offer contributions to research on implementation on the high-stakes testing components on NCLB through participation in the study. Principals have been rich sources of information for a variety of studies as they are involved in the efficient and effective implementation of education reform.

CHAPTER 4

REPORT OF DATA AND DATA ANALYSIS

This chapter includes a brief introduction to the study including the purpose of the study and a summary of the research methodology. Also included are the research questions to be answered, a factual reporting of the data gathered, and an interpretation of the data. This chapter includes the researcher's answers to the research questions based on her findings in the study.

Successful implementation of high-stakes testing within the current national, state, and local directives has been attributed to the principal as leader of the school organization. Georgia elementary principals began the task of providing a quality education for every student by the 2013-2014 school year with the guarantee that no student would be left behind according to the No Child Left Behind Act of 2001 (NCLB). Accountability for this performance goal is based on student performance on high-stakes testing. Specific components of the reform pertinent to elementary schools included emphasized evaluation of student performance through high-stakes assessments, the provision of standards with detailed testing plans designed to determine accomplishment of the standards, and data-driven research-informed instruction.

Introduction

The study was designed to explore the perception of Georgia public elementary principals' regarding utilization of high-stakes testing as the criteria for the determination of the accomplishments of public elementary school principals. The survey was self-designed and consisted of three major sections: 1) personal and demographic information, 2) principals' general perceptions of the utilization of high-stakes testing

as the determinant of student achievement, and 3) two open-ended questions eliciting opinions about the influence of high-stakes testing on principal accountability for student achievement and actual student achievement.

The survey was examined by an expert panel to determine face validity, and was pilot tested by 10 elementary principals in Alabama to determine content validity and reliability. Suggestions offered by the expert panel enabled the researcher to clarify questions, and provided assistance with design of the survey. Responses from pilot participants indicated that no further revisions were necessary. The survey was mailed to 610 Georgia public elementary principals randomly selected from the Georgia Public School Directory published by the Georgia State Department of Education (2006), after the revisions were completed.

Data analysis was conducted utilizing the SPSS (SPSS, 2003). Data analysis utilizing SPSS generated descriptive statistics such as frequencies, means, and standard deviations. The two open-ended questions were used to obtain a narrative explanation of principals' perceptions regarding the perceived influence of high-stakes testing on the accountability of the principals for student achievement, and the impact of high-stakes testing on student achievement. Open-ended questions were analyzed by searching for patterns, themes and categories (Creswell, 2003).

Research Questions

The major research question was, "What are the perceptions of Georgia elementary principals about high-stakes testing?" The following sub-questions guided the research in determining principals' perceptions of the purpose of high-stakes testing, improvement

of academic performance of students, measurement of the effectiveness of principals, and the use of rewards and sanctions. The study investigated five research questions:

1. What do Georgia elementary principals perceive the purpose of high-stakes testing to be?
2. Do Georgia elementary principals perceive that high-stakes testing improves academic performance?
3. Do Georgia elementary principals perceive that high-stakes testing appropriately measures the effectiveness of their performance?
4. What are the perceptions of Georgia elementary principals regarding rewards and sanctions for high-stakes testing?
5. Do demographic variables make any difference in principals' perceptions of high-stakes testing?

Findings

Principals' Personal and Professional Demographics

A total of 335 Georgia elementary principals responded to the survey. The 335 responses represented a 55% return rate of the total mailing.

The current status of Georgia elementary principals was obtained by analyzing demographic data in six questions of the survey. Table 2 presented data from these elementary principals regarding their years of experience as principals.

Respondents revealed that they had served from 1 to 26+ years in the principalship. The standard deviation indicated a low amount of variability in years served as principals among respondents, and the average length of time spent in the principalship was reported as two years. The largest percentage of respondents revealed that they had the

least experience as principals with 46.9% reporting from 1 to 5 years of experience.

Approximately 25% indicated that they had from 6 to 10 years of experience, while less than 5% indicated that they had remained in the principalship for 16 – 20 years, and less than 5% indicated that they had remained in the principalship for 21-25 years.

The most frequently reported number of years' experience of those participating in the study was found to be 1-5 years (46.9%).

Further exploration of respondents' characteristics was accomplished by examining the AYP status of schools. The 2005-06 AYP status of the principals, separated into two categories was reported in Table 3. The value of "1" was assigned to the "met" category, and "2" to the "did not meet" category. Most of the principals indicated that they made AYP for the 2005-06 year, with 91.6% making AYP and 8.4% not making AYP.

Table 2

Respondents' Years of Experience as Principal

Variable	Category	Percentage	<u>n</u>
Years' Experience			335
1-5	1	46.9%	
6-10	2	24.5%	
11-15	3	9.9%	
16-20	4	10.1%	
21-25	5	4.5%	
26+	6	4.1%	

Table 3

Principals' 2005-06 AYP Status

Variable	Category	Percentage	<u>n</u>
AYP Status			335
Met	1	91.6%	
Did Not Meet	2	8.4%	

Table 4 reported principals' age, separated into four categories. The value of "1" was assigned to the "under 35" category, "2" to the "35-45" category, "3" to the "46-55" category, and "4" to the "56+" category.

Table 4

Principals' Age Reported in Categories

Variable	Category	Percentage	<u>n</u>
Age			335
Under 35	1	6.3%	
35-45	2	28.4%	
46-55	3	45.7%	
56+	4	19.6%	

The age-range most frequently reported was represented by the “46-55” category, with almost 45.7% of the study population occupying this age-range. The category representing the under 35 age-range comprised 6.3% of those reporting. The second most frequently reported age-range was revealed as the “35-45” category, with 28.4% of the respondents revealing this information. The 56+ age range was represented by 19.7% of the responding principals.

Gender results have been reflected in Table 5. The value of “1” was assigned to the “male” category, and “2” was assigned to the “female” category.

Table 5

Principals’ Gender

Variable	Category	Percentage	<u>n</u>
Gender			335
Male	1	35.5%	
Female	2	64.5%	

More females than males responded to the survey with almost two thirds of the respondent population comprising the female gender. Georgia female elementary principals were represented by 64.2% and male respondents encompassed 35.5% of those responding.

Principals' educational level was divided into three categories with values of "1", "2", and "3" assigned to the Master's, Educational Specialist, and Doctorate degrees respectively. Table 6 reported these data.

The highest educational level reported by most principals was the Education Specialist degree, with 60.9%, revealing that they held this credential. Slightly more than 25% of the respondents reported holding the Doctorate degree, and 12.5% of the respondents revealed that they held the Master's degree.

Table 6

Principals' Educational Level

Variable	Category	Percentage	<u>n</u>
Educational Level			335
Master's	1	12.5%	
Education Specialist	2	60.9%	
Doctorate	3	26.6%	

Table 7 reported the data collected regarding school communities served by the respondents. Values of "1", "2", and "3" were assigned to represent the communities with "1" representing the urban area, "2" representing suburban areas, and "3" representing the rural areas.

Table 7

School Community Served

Variable	Category	Percentage	<u>n</u>
School Community			335
Urban	1	16.1%	
Suburban	2	44.8%	
Rural	3	39.1%	

Most of the respondents reported working in suburban communities (44.2%), however, this representation only marginally surpassed those who reported working in rural areas (39.1%). Urban communities were represented by 16.1% of the respondents.

Principals' Perceptions

Thirteen survey statements were posed to assess principals' beliefs regarding the overarching implications addressed by high-stakes testing, and to answer the research questions: 1) What do Georgia elementary principals perceive the purpose of high-stakes testing to be? 2) Do Georgia elementary principals perceive that high-stakes testing improves academic performance? 3) Do Georgia elementary principals perceive that high-stakes testing appropriately measures the effectiveness of their performance? 4) What are the perceptions of Georgia elementary principals regarding rewards and sanctions for high-stakes testing? Each statement presented was aligned with responsibility issues principals will face as a result of this implementation. Values ranging from "1" representing "strongly disagree" to "4" representing "strongly agree" were provided as response choices.

Table 8 provided frequencies and descriptive statistics of elementary principals' responses to the first question in the second section of the survey, referring to the accountability of the principal for student achievement as measured by high-stakes testing.

Table 8

Principals Should be held Accountable for Student Achievement as Measured by High-Stakes Testing

Variable	Category	Frequency	Percentage	<u>M</u>	<u>SD</u>	<u>n</u>
Accountability				2.91	.79	335
Strongly Agree	4	70	20.9%			
Agree	3	184	54.9%			
Disagree	2	58	17.3%			
Strongly Disagree	1	21	6.3%			
Did Not Respond	0	2	0.6%			

Three fourths of the respondents (75.8%) agreed or strongly agreed that they should be held accountable for student achievement as measured by high-stakes testing. One fourth (23.6%) did not believe they should be accountable based on high-stakes testing results. The results include a mean of 2.91 and a standard deviation of .79.

The second survey question within the second section requested principals to indicate if student achievement would improve as the school received rewards based on high-stakes testing. Table 9 revealed their responses descriptively.

Table 9

*Student Achievement will Improve as Schools Receive Rewards Based on High-Stakes**Testing*

Variable	Category	Frequency	Percentage	<u>M</u>	<u>SD</u>	<u>n</u>
Rewards				2.38	.82	335
Strongly Agree	4	25	7.5%			
Agree	3	123	36.7%			
Disagree	2	139	41.5%			
Strongly Disagree	1	46	13.7%			
Did Not Respond	0	2	0.6%			

Considerable agreement (44.2%) that student achievement would improve if their school received rewards based on high-stakes testing was revealed. Slightly more than one half (56.3%) of the principals disagreed that student achievement at their school would improve with the receipt of rewards based on high-stakes testing. The results include a mean of 2.38 and a standard deviation of .82.

The third survey question in the second section explored principals' beliefs regarding whether student achievement would improve if their school received sanctions based on high-stakes testing. These data are revealed in Table 10.

Table 10

Student Achievement will Improve as Schools Receive Sanctions Based on High-Stakes

Testing

Variable	Category	Frequency	Percentage	<u>M</u>	<u>SD</u>	<u>n</u>
Sanctions				2.11	.86	335
Strongly Agree	4	18	5.4%			
Agree	3	88	26.3%			
Disagree	2	140	41.7%			
Strongly Disagree	1	87	26.0%			
Did Not Respond	0	2	0.6%			

One third (31.7%) of the principals agreed that student achievement would improve if the school received sanctions based on high-stakes testing. A large percentage (67.8%) did not believe that student achievement would improve if their school received sanctions based on high-stakes testing. The results include a mean of 2.11 and a standard deviation of .86.

The fourth question requested principals to provide their opinions regarding whether or not high-stakes testing was fair. These data are reported descriptively in Table 11.

Table 11

High-Stakes Testing is Fair

Variable	Category	Frequency	Percentage	<u>M</u>	<u>SD</u>	<u>n</u>
Fairness				2.13	.81	335
Strongly Agree	4	13	3.9%			
Agree	3	93	27.8%			
Disagree	2	149	44.5%			
Strongly Disagree	1	77	23.0%			
Did Not Respond	0	3	0.8%			

Slightly less than one third (31.7%) of the principals believed that high-stakes testing was fair. A larger percentage (67.5%) did not believe that high-stakes testing was fair. The results include a mean of 2.13 and a standard deviation of .81.

The fifth survey question in the second section explored principals' beliefs regarding whether they perceived high-stakes testing would improve the education of all students. These data are revealed descriptively in Table 12.

Table 12

High-Stakes Testing Improves Education for all Students

Variable	Category	Frequency	Percentage	<u>M</u>	<u>SD</u>	<u>n</u>
Improvement				2.27	.81	335
Strongly Agree	4	18	5.4%			
Agree	3	110	32.8%			
Disagree	2	148	44.2%			
Strongly Disagree	1	57	17.0%			
Did Not Respond	0	2	0.6%			

More than one third of the elementary principals (38.2%) perceived that high-stakes testing had improved education for all students. However, 61.2% of the principals did not perceive that high-stakes testing had improved education for all students with 17% strongly disagreeing with improvements. The results include a mean of 2.27 and a standard deviation of .81.

The sixth survey question explored principals' beliefs regarding whether they perceived that the high-stakes testing results of their school were an appropriate measure of their accountability. These data are reported descriptively in Table 13.

Table 13

High-Stakes Testing Results Appropriately Measures Accountability of Principal

Variable	Category	Frequency	Percentage	<u>M</u>	<u>SD</u>	<u>n</u>
Accountability				2.34	.75	335
Strongly Agree	4	13	3.9%			
Agree	3	128	38.2%			
Disagree	2	149	44.5%			
Strongly Disagree	1	42	12.5%			
Did Not Respond	0	3	0.9%			

Data indicated agreement (42.1%) that high-stakes testing was an appropriate measure of principal accountability, with a small percentage (3.9%) strongly agreeing. When the “strongly disagree” and “disagree” ratings were combined, 57.% of the respondents negatively viewed high-stakes testing as an effective measure of the accountability of the principal. The results include a mean of 2.34 and a standard deviation of .75.

The seventh survey question requested principals’ to assess their use of high-stakes testing results to improve student achievement. These data are revealed descriptively in Table 14.

Table 14

Principals Consistently Use High-Stakes Testing Results to Improve Student Achievement

Variable	Category	Frequency	Percentage	<u>M</u>	<u>SD</u>	<u>n</u>
Achievement				3.02	.75	335
Strongly Agree	4	84	25.1%			
Agree	3	187	55.8%			
Disagree	2	50	14.9%			
Strongly Disagree	1	13	3.9%			
Did Not Respond	0	1	0.3%			

Elementary principals overwhelmingly believed (80.9%) in the consistent utilization of high-stakes testing results to improve student achievement. A moderate percentage (18.8%) of the respondents revealed uncertainty with respect to the use of high-stakes testing for improvement of student achievement. The results include a mean of 3.02 and a standard deviation of .75.

The eighth survey question requested principals to provide their opinions regarding whether or not factors beyond their control influences student achievement and should be considered with high-stakes testing. These data are reported descriptively in Table 15.

Table 15

Factors beyond Principal Control Influence High-Stakes Testing and Should be Considered

Variable	Category	Frequency	Percentage	<u>M</u>	<u>SD</u>	<u>n</u>
Other Factors				3.33	.82	335
Strongly Agree	4	172	51.3%			
Agree	3	113	33.7%			
Disagree	2	34	10.1%			
Strongly Disagree	1	14	4.3%			
Did Not Respond	0	2	0.6%			

An overwhelming 85% of the principals believed that factors beyond the control of the principal impacted student achievement and should be considered with high-stakes testing. A moderate percentage (14.3%) of the respondents did not believe that factors beyond their control impacted student achievement and should be considered. The results include a mean of 3.33 and a standard deviation of .82.

The ninth survey question explored principals' beliefs regarding whether they perceived the purpose of high-stakes testing was to improve instructional leadership. These data are revealed in Table 16.

Table 16

*Purpose of High-Stakes Testing is to Make Principals More Focused Instructional**Leaders*

Variable	Category	Frequency	Percentage	<u>M</u>	<u>SD</u>	<u>n</u>
Focus				2.42	.94	335
Strongly Agree	4	42	12.5%			
Agree	3	121	36.1%			
Disagree	2	105	31.3%			
Strongly Disagree	1	64	19.2%			
Did Not Respond	0	3	0.9%			

Considerable agreement (48.6%) that the purpose of high-stakes testing was to improve the instructional leadership of the principal was revealed. One half (50.4%) of the respondents disagreed or strongly disagreed that the purpose was to improve the principal's instructional leadership focus. The standard deviation statistic indicated that principals' responses varied to a great extent on this item than on any other quantitatively analyzed item on the survey. The results include a mean of 2.42 and a standard deviation of .94.

The tenth survey question requested principals to provide their opinions regarding whether or not their school's high-stakes testing results accurately evaluated their school leadership. These data are reported descriptively in Table 17.

Table 17

High-Stakes Testing Results of School Accurately Evaluate Principals' School Leadership

Variable	Category	Frequency	Percentage	<u>M</u>	<u>SD</u>	<u>n</u>
Focus				2.25	.76	335
Strongly Agree	4	14	4.2%			
Agree	3	106	31.6%			
Disagree	2	161	48.1%			
Strongly Disagree	1	50	14.9%			
Did Not Respond	0	4	1.2%			

One third of the respondents (35.8%) believed that the high-stakes testing results of the school accurately evaluated the school leadership of the principal. Almost two thirds (63%) did not believe that the school leadership of the principal was accurately evaluated by the high-stakes testing results. The results include a mean of 2.25 and a standard deviation of .76.

The eleventh survey question requested principals to provide their opinions whether or not the high-stakes testing results of their school determined their perception of high-stakes testing. These data are reported descriptively in Table 18.

Table 18

High-Stakes Testing Results Determine Principals' Perceptions of High-Stakes Testing

Variable	Category	Frequency	Percentage	<u>M</u>	<u>SD</u>	<u>n</u>
Focus				2.07	.79	335
Strongly Agree	4	9	2.7%			
Agree	3	91	27.2%			
Disagree	2	151	45.1%			
Strongly Disagree	1	83	24.7%			
Did Not Respond	0	1	0.3%			

Almost one third (29.9%) of the respondents agreed that the test results of their school determined their perception of high-stakes testing. Almost 70% of the principals indicated that their perception of high-stakes testing was based on factor other than the results from their school. The results include a mean of 2.07 and a standard deviation of .79.

The twelfth survey question explored principals' beliefs regarding whether they perceived the purpose of high-stakes testing was to improve student achievement. These data are revealed in Table 19.

Table 19

Purpose of High-Stakes Testing is to Improve Student Achievement

Variable	Category	Frequency	Percentage	<u>M</u>	<u>SD</u>	<u>n</u>
Focus				2.74	.84	335
Strongly Agree	4	57	17.0%			
Agree	3	160	47.8%			
Disagree	2	87	26.5%			
Strongly Disagree	1	28	8.4%			
Did Not Respond	0	1	0.3%			

Two thirds of the respondents (64.8%) agreed or strongly agreed that the purpose of high-stakes testing was to improve student achievement. A moderate percentage (34.5%) did not believe that the improvement of student achievement was the purpose of high-stakes testing. The results include a mean of 2.74 and a standard deviation of .84.

The last survey question requested principals to provide their opinion regarding whether or not the purpose of high-stakes testing was to hold principals accountable for their performance. These data are reported descriptively in Table 20.

Table 20

Purpose of High-Stakes Testing is to Hold Principals' Accountable for Performance

Variable	Category	Frequency	Percentage	<u>M</u>	<u>SD</u>	<u>n</u>
Focus				2.56	.86	335
Strongly Agree	4	38	11.3%			
Agree	3	150	44.8%			
Disagree	2	101	30.1%			
Strongly Disagree	1	42	12.5%			
Did Not Respond	0	4	1.3%			

Considerable agreement (56.1%) that the purpose of high-stakes testing was to hold principals accountable for their performance was revealed, while 42.6% of the principals did not perceive the purpose of high-stakes testing to be principal accountability. The results include a mean of 2.56 and a standard deviation of .86.

Principals' Open-Ended Responses

Two open-ended questions were posted in qualitative format to answer the second and third research questions that explored the perceptions of high-stakes testing. These questions were designed to assess whether or not principals believed high-stakes testing had influenced the way they held themselves accountable for student achievement or if high-stakes testing had impacted student achievement. Textual responses were examined for establishment of themes. Data has been presented in numerical format according to categorical responses in Table 21.

Table 21

Open-Ended Question 1 (Survey Question 14) How will high-stakes testing influence the way you hold yourself accountable for student achievement in your school?

Category of Responses	Number of Respondents	<u>n</u>
		250
Data driven/test results decisions	106	
No influence, negative influence, increased stress	66	
Increased focus on student achievement	34	
Increased focus on instructional strategies	20	
More accountability	18	
Increased attention for subgroups	16	
Use test results as only one indicator	15	
Standards driven accountability	11	
High expectations for all students	11	

In response to question 14, “How will high-stakes testing influence the way you hold yourself accountable for student achievement in your school?” principals revealed that they would make more data driven decisions. A number of patterns emerged through data analysis that supported this contention.

Principals noted that data was analyzed to identify individual, school, and systemic strengths and weaknesses to determine the appropriate resources, professional learning opportunities, instructional strategies, and budget demands to improve student learning.

Succinctly summarizing this concept, “The analysis of the data will be used to determine student needs. It is then my job to make sure teachers have the knowledge and resources they need to improve student learning” was reported by one respondent (Respondent 73). Echoing this sentiment, another respondent revealed “High-stakes testing influences the way I hold myself accountable in several areas: providing professional development and resources, hiring the best and the brightest, scheduling students, monitoring and analyzing data to guide instruction” (Respondent 133).

Despite the positive regard for educational decisions based on an analysis of the high-stakes testing results, principals often reported a continuum of negative responses from the lack of any influence to negative influence to significant increased levels of stress. One respondent wrote “We have always held ourselves accountable for all students learning to the best of their abilities. High-stakes testing has caused teachers to push harder when it was not needed or necessary, and has caused teachers to question their value as educators in a system that relies on high-stakes testing only” (Respondent 9). “I will always feel accountable for my school’s performance. Testing just adds another and higher dimension of stress. It is particularly difficult when we have factors out of the school’s control” (Respondent 130). One respondent wrote “I’m not sure it will change me at all. You can lead a horse to water but cannot make him drink” (Respondent 1). Respondent 302 reported, “My school does well on high-stakes testing but at what a cost! It is only one part/aspect of my accountability. Day to day performance of students and staff are of a greater concern to me as well as a love of learning and teaching. Creating an atmosphere that promotes this love will lead to real learning.”

Increased focus on student achievement was revealed as an influence of high-stakes testing on the way principals held themselves accountable for student achievement in their school. One respondent (87) reported, “Student achievement is kept on the front burner – all decisions must be filtered through the student achievement filter” objectively reflecting on the effects of high-stakes testing on accountability. “High-stakes testing influences programs necessary for greater student achievement. I must make the right choices for my students” (Respondent 96) and “Whatever the outcome is, I will take a personal stand to increase my knowledge to move students forward” (Respondent 99) were examples of the respect for yet anxiety reflected by principals relating to improving student achievement.

A number of principals indicated that they would increase their focus on the instructional opportunities utilized for instruction. “It is imperative that educators continuously work to improve the learning opportunities for all children” (Respondent 57), “I will use data from the tests to improve instruction” (Respondent 69), “It makes you become more focused on instruction” (Respondent 142) were examples of the influence on their accountability.

Some principals reported that the influence of high-stakes testing was to increase the way they held themselves accountable for student achievement in their school. “High-stakes testing is at the top of my accountability list for school success” (Respondent 173), “I will hold myself accountable if our student achievement results decrease and will seek input from all stakeholders to bring about improvement” (Respondent 189) and “I feel accountable for student achievement and I will make changes to ensure improvement” (Respondent 492) were examples of their perceptions.

Increased attention for subgroups was reported as a resulting influence on principal accountability. Summarizing this concept, “I will focus on subgroups” (Respondent 240), “I will focus on analyzing subgroups” (261), and “I will improve all subgroups” (Respondent 311). Echoing this sentiment, another respondent reported “To make AYP I must make sure every subgroup is succeeding” (Respondent 442).

Factored in with the expressed positive influence of high-stakes testing, some principals emphasized the fact that they utilized it as only one source of information. Statements such as “I use many assessments to guide instruction, I never depend on one assessment (CRCT)” (Respondent 280), and “I will continue to use multiple sources of information (demographics, ITBS scores, local assessments) to inform instructional practices at my school” (Respondent 379) were examples of their perceptions.

A number of individuals indicated that standards driven accountability and high expectations for all students were resultant influences on their accountability. Of those responding to this issues, a number related that they were focused on ensuring that students were engaged in authentic learning experiences in standards-based classrooms and believing that 100% of students could meet standards given the appropriate instruction.

The second research question was explored further by survey question 15 that asked, “How has high-stakes testing impacted student achievement in your school?” and required a written response. Table 22 revealed numerical totals for individuals’ responses to specified categories.

Table 22

Open-Ended Question 2 (Survey Question 15) How has high-stakes testing impacted student achievement in your school?

Category of Responses	Number of Respondents	<u>n</u>
Positive impact	190	272
Negative impact	46	
Little or neutral impact	37	

Overwhelmingly, principals were supportive and positive toward the impact of high-stakes testing in their school. A number of patterns emerged through data analysis that supported this contention.

Principals noted that analysis of the test data was paramount for more focused and differentiated instruction for increased student achievement. Statements such as “I look at individual student scores and make decisions based on those scores for what needs to be done to help guide instruction” (Respondent 18), “We have developed a data-driven laser-like focus in school improvement” (Respondent 42), “We really teach our teachers how to analyze data and use that data to improve students’ strengths and weaknesses” (Respondent 47), and “Our scores are good but we have to look at the data each year to see how we are doing because the outcomes are so critical” (Respondent 69) were examples of the utilizations reflected by principals relating to data analysis.

Of those principals indicating that high-stakes testing had made them much more aware of subgroups and the need to focus on their strengths and weaknesses, more principals appeared concerned with the fact that students with disabilities or LEP students could keep a school from making AYP. Frustration with this component was exemplified by the comment from Respondent 159, who stated “I really feel for our special education students. I think special education and GAA (Georgia Alternative Assessment) should not be able to place a school on AYP. I am okay with testing students overall to make sure standards are being addressed but special education students should not be able to cause a school to be in Needs Improvement”.

The added high-stakes testing responsibilities mandated by NCLB have been perceived to have little or no impact by some principals. Of those who responded that this was a concern, a number indicated concern about a correlation between high-stakes testing and achievement. Remarks such as “If a correlation between high-stakes testing and achievement has been done at this school or any other schools for that matter, I’m not aware of it (Respondent 2) and “Test results might suggest that student achievement has increased however, there is some question as to if those gains are meaningful (Respondent 389) served to illustrate this apprehension.

Although not overwhelming, a number of respondents alluded to the stress and pressure experienced as a results of the high-stakes testing mandates. Respondent 464 reported, “Teachers are stressed, kids are stressed, it discourages high level teaching and learning” and Respondent 313 reported, “It results in a great deal of stress”.

The researcher intended to discover whether the personal and professional information gathered on Georgia elementary principals could be utilized as predictors of

their responses to the 13 survey questions as indicated in research question, 5. Do demographic variables make any difference in principal's perceptions of high-stakes testing?

The one-way analysis of variance statistical procedure was employed to examine differences in principals' perceptions regarding high-stakes testing according to (a) length of service; (b) AYP status for 2005-06 school year; (c) age; (d) gender; (e) educational level (M. Ed., Ed., S., Ed. D./Ph. D.) or (f) school geographical location (urban, suburban, rural)?

Analysis of the data revealed no significant differences among length of service, age, gender, educational level, and school community served when responses to the survey questions were compared. Tables 23, 24, 25, 26, and 27 reported the results of these data.

Table 23

ANOVA Summary Table for Length of Service and Survey Questions

Source	SS	df	MS	F	Sig.
Between Groups	1.224	5	.245	1.047	.390
Within Groups	72.910	312	.234		
Total	74.134				

Table 24

ANOVA Summary Table for Age and Survey Questions

Source	SS	df	MS	F	Sig.
Between Groups	1.537	3	.512	2.216	.086
Within Groups	72.597	314	.231		
Total	74.134	317			

Table 25

ANOVA Summary Table for Gender and Survey Questions

Source	SS	df	MS	F	Sig.
Between Groups	.812	1	.812	3.501	.062
Within Groups	73.321	316	.232		
Total	74.134	317			

Table 26

ANOVA Summary Table for Educational Level and Survey Questions

Source	SS	df	MS	F	Sig.
Between Groups	.587	2	.293	1.256	.286
Within Groups	73.547	315	.233		
Total	74.134	317			

Table 27

ANOVA Summary Table for School Community Served and Survey Questions

Source	SS	df	MS	F	Sig.
Between Groups	.124	2	.062	.263	.769
Within Groups	74.010	315	.235		
Total	74.134	317			

Analysis of the data examining whether AYP status for the 2005-06 school year had any effect on principals' perceptions revealed significant differences among the groups on the survey questions. Table 28 reported results of the comparison of principals

meeting or not meeting AYP status during the 2005-06 school year regarding perceptions of high-stakes testing.

The mean for the group meeting AYP status for the 2005-06 school year was 2.53 and the mean for the group not making AYP status for the 2005-06 school year was 2.25.

Table 28

ANOVA Summary Table for AYP Status for 2005-06 School Year and Survey Questions

Source	SS	df	M	F	Sig.
Between Groups	1.981	1	2.53	8.675	.003
Within Groups	72.153	316	2.25		
Total	74.134	317			

Summary

Georgia elementary principals who responded were primarily 46-55 year-old females who worked in suburban areas of the state. Typically, they possessed the education specialist degree and had an average of two years experience in their positions.

These individuals understood the high-stakes components of NCLB, possessed skills to use test results to improve student achievement, and supported principal accountability for student achievement as measured by high-stakes testing. However, they strongly supported consideration of the factors beyond their control that influenced student achievement.

Principals agreed that the purpose of high-stakes testing was to improve student achievement but that high-stakes testing was not fair and did not improve education for all students. Principals perceived that high-stakes testing did not improve student achievement when schools received sanctions based on high-stakes testing results, or evaluate the school leadership of principals. Principals also agreed that their perceptions of high-stakes testing were not based on their test scores.

Principals did not agree that student achievement improved as schools received rewards based on testing results or that high-stakes testing results were appropriate measures of principal accountability. They did not agree that the purpose of high-stakes testing was to make principals more focused instructional leaders or to hold them accountable for performance.

CHAPTER 5

SUMMARY, CONCLUSIONS, AND IMPLICATIONS

This chapter has been designed as a synopsis of the total study and has been organized through a brief introduction of the study followed by the research findings of the study and a discussion of these findings as they related to the comparable data found in the literature review. The conclusions, implications, and recommendations drawn from the research findings follow the discussion.

Introduction

What are Georgia elementary principals' perceptions of high-stakes testing? The intent of this study was to answer this question through the development of a theoretical structure as the viewpoints of the participants became known. Specific sub-questions were used by the researcher to answer the above over-arching question. These sub-questions were:

1. What do Georgia elementary principals perceive the purpose of high-stakes testing to be?
2. Do Georgia elementary principals perceive that high-stakes testing improves academic performance?
3. Do Georgia elementary principals perceive that high-stakes testing appropriately measures the effectiveness of their performance?
4. What are the perceptions of Georgia elementary principals regarding rewards and sanctions for high-stakes testing?
5. Do demographic variables may any difference in principal's perceptions of high-stakes testing?

Principals, as the leaders of their schools, are compelled by the responsibility of their positions to implement local, state, and national mandates in an effective and efficient manner (Finch, 1999). It is necessary for principals to provide an atmosphere where reform implementation can be successful (Fullan, 2005).

The responsibility of the principal during the implementation of reform has not been defined by educational researchers. It was the intention of the researcher to request information from a large number of Georgia elementary principals regarding high-stakes testing perceptions during implementation of the No Child Left Behind Act of 2001 in Georgia elementary public schools. No research previously conducted could be located that explored Georgia principals' perceptions of high-stakes testing.

The descriptive study design was utilized to establish the current status of Georgia elementary principals. It was intended to determine differences among principals from different areas of the state, from different educational levels, ages, gender, experience, and according to 2005-06 AYP status. The study provided a cross-sectional sample of Georgia elementary principals for their perceptions of high-stakes testing.

The study was conducted in three phases. The first phase consisted of selecting an expert panel for the purpose of establishing face validity of the self-designed survey. The second phase included a pilot survey conducted with 10 Alabama elementary principals after face validity was established, to satisfy validity and reliability requirements and to further refine the intelligibility of the survey. The third phase consisted of conducting the study with randomly sampled with replacement selected principals from Georgia. Both quantitative and qualitative questions were included in the survey in order to explore principals' perceptions objectively and thematically.

To ensure that the prescribed sample size of 302 (Krejcie & Morgan, 1970) was obtained, 610 surveys were mailed to elementary principals in Georgia. Since 1,276 elementary principals have been currently designated as elementary principals in Georgia, (Georgia Public Schools Directory, Georgia Department of Education, 2006), the sample size used represented a universal population of 1,276. A second complete survey mailing was conducted for the purpose of contacting nonrespondents. A total of 335 surveys were received, representing 55% of the total mailing.

Research Findings

In regard to the impact of demographic variables on principal's perceptions of high-stakes testing, Georgia principals responding to the survey were typically females between the ages of 46 and 55. They worked predominately in suburban communities, and had between 1 and 5 years of experience in their positions. Although the suburban communities were the most frequently reported location within which principals worked, rural communities were represented by almost as many principals. A majority of those individuals held the education specialist degree, and an overwhelmingly majority (91.6%) met AYP for the 2005-06 school year. The demographic variable with statistical significance was whether the principal made AYP for the 2005-06 school year.

Georgia elementary principals perceived that high-stakes testing improved academic performance. However, principals perceived that factors beyond their control influenced student achievement and should also be considered with the high-stakes testing results of schools (85%). The perception was widely held that the consistent use of high-stakes testing results improved student achievement (81%). Principals perceived

that they should be held accountable for student achievement as measured by high-stakes testing (76%).

Inquiring as to principals' perceptions of the purpose of high-stakes testing, principals perceived that the purpose of high-stakes testing was to improve student achievement in their schools (65%). More than half (56%) of the principals believe that the purpose of high-stakes testing was to hold principals accountable for their performance and half (49%) believed the purpose was to make principals better focused instructional leaders.

Principals perceived that factors other than the actual high-stakes testing results of their schools determined their perceptions of high-stakes testing (70%), evaluation of their school leadership (63%), and measurement of their accountability (57%). Principals did not perceive that high-stakes testing appropriately measured the effectiveness of their performance as principals.

With regard to principals' beliefs about the fairness of high-stakes testing, more than two thirds (68%) believe high-stakes testing was not fair, with 23% strongly holding this belief. The perception was widely held that high-stakes testing did not improve education for all students (61%). Half (55%) of the principals believed that student achievement would not improve with the receipt of school rewards based on high-stakes testing and two thirds (68%) agreed that student achievement would not improve with the receipt of school sanctions based on high-stakes testing.

Discussion of the Research Findings

Principals' perceptions of high-stakes testing, principals' personal and professional characteristics, the influence of high-stakes testing upon principals' personal

accountability for student achievement and the impact of high-stakes testing upon student achievement were discussed in relation to the findings and data analysis, and review of the literature. The study endeavored to determine whether Georgia elementary principals believed that the purpose of high-stakes testing was to make principals more focused instructional leaders, or to improve student achievement, or to hold principals accountable for their performance, if high-stakes testing had actually improved student achievement or appropriately measured the effectiveness of principals, the impact of rewards and sanctions for high-stakes testing, and finally, whether demographic data affected the manner in which they responded to specific issues.

Policymakers have claimed that the purpose of high-stakes testing was to guarantee parents that students passing the tests have mastered the skills measure by the tests (Popham, 2001; Ohanian, 1999). In actuality, legislators and educational policymakers were looking at schools that were assumed to be falling down on the job of teaching since children were making their way through the educational system without having learned how to read, write or complete job applications (Popham, 2001; Sclafani, 2002). The study provided evidence that many Georgia elementary principals perceived that the purpose of high-stakes testing was to improve student achievement in schools but they did not agree that the purpose was to hold principals accountable for their performance, or to make them more focused instructional leaders.

School reform mandates have nurtured the notion that students' test scores could determine the quality of a school. Significant amounts of money have been dispensed to school districts to support programs designed to improve student achievement. With increased funding, corresponding safeguards in the form of annually administered

high-stakes testing have been built in to monitor the federal funds (Popham, 2001). Georgia elementary principals characterized themselves as accountable for student achievement as measured by high-stakes tests and as improvers of student achievement with the consistent use of high-stakes testing results. However elementary principals in Georgia strongly perceived that factors beyond their control that influence student achievement should also be considered with the high-stakes testing results of schools.

A bipartisan effort, NCLB represented the greatest federal incursion into K-12 education to date (Gallup & Rose, 2002). The plan called for increased federal attention to standards, assessment and accountability, and implementation of federal government priorities with federal rewards and sanctions based on meeting acceptable levels of educational change (Education Commission of the States, 2001). Many principals communicated that student achievement did not improve when schools received sanctions based on high-stakes testing results and there was considerable disagreement that student achievement improved when schools received rewards based on the results.

Examination of principals' years of experience, age, gender, educational level, and school community served revealed no significant differences among the groups when their responses to the survey questions were compared. Analysis of the data examining whether AYP status for the 2005-06 school year had any effect on principals' perceptions revealed significant differences among the groups.

Conclusions

The following conclusions were drawn from the analysis of the research findings. They represent the researcher's interpretation of this analysis of findings:

1. Georgia elementary principals responding to this survey can be characterized

as primarily females, between the ages of 46-55 who worked in suburban communities. They held the education specialist degree, averaged 2 years experience as principals, and met AYP status for the 2005-06 school year. The principals were young with few years of experience and tended to have positive regard for the law. They are also currently making AYP so these positive results at this time may influence their opinions.

2. The high-stakes testing components of NCLB, a law that stems from the basic assumption that every child – regardless of income, gender, race, ethnicity, or disability, can learn and deserves to learn has been understood by the majority of principals. These individuals expressed confidence in their abilities to utilize high-stakes testing results to improve student achievement and to be held accountable for student achievement. At the same time, these principals expressed concern that factors beyond their control influenced student achievement and should be considered with test results. Principals expressed positive attitudes about high-stakes testing and at the same time expressed strong concerns about the use of high-stakes testing.
3. The purpose of high-stakes testing is the improvement of student achievement in schools. Again, principals expressed respect for the current mandates but they also expressed concerns about the negative impact of the testing upon schools.
4. Although the high-stakes tests are consistently used to improve student achievement, they do not improve education for all students. Principals

expressed concerns about special education students and students below grade level, the students NCLB mandated saving.

5. High-stakes testing appropriately holds principals accountable for student achievement as measured by the tests, but does not accurately evaluate their school leadership or accountability as a principal.
6. Student achievement improves as schools receive rewards based on high-stakes testing but not with the receipt of sanctions based on the results.
7. Georgia elementary principals' perceptions of high-stakes testing did not vary according to length of experience, age, gender, educational level, or school communities served, however the status of AYP for the 2005-06 school year appeared to have some bearing on their perceptions. Principals meeting AYP status for the 2005-06 school year reported more positive perceptions.

Implications

The study should inform the policy makers of Georgia of the importance of receiving input from the elementary principals of Georgia for efforts to address comprehensive education reform. The principals presently serving as leaders of schools have valuable insights into the demands to benefit all students and subgroups of students with meaningful high standards, at the required Adequate Yearly Progress, AYP pace. The demands create unique situations that require more than a "one size fits all" approach to school improvement. The principal is pivotal for leading the school in the provision of a quality education for every student by the 2013-2014 school year and inclusion in decision-making has serious implications for successful implementation of school reform.

The utilization of high-stakes testing continues to exist as a controversial issue in the theme of accountability and the results of this study support that finding. An accurate inventory of the current assessments that are already working in public schools within Georgia could well serve policy makers with accountability measures for school improvement. There is no need for a new law to design new tests to measure new curriculum which results in huge amounts of time and money being diverted away from the current task at hand – to improve student achievement of all children.

It is imperative that policymakers examine the demands of coercive accountability, and the use of standardized tests as the sole criteria for determining the progress of public school principals, teachers, and students. Georgia elementary principals indicated professionalism in attempting to be held accountable for student achievement based on high-stakes testing results. Despite this professionalism, principals revealed frustration with the unfairness of the tests the lack of consideration for the factors beyond the control of the principal that influenced the results.

The study should serve to inform policy makers, the Georgia State Department of Education, and local Boards of Education that the use of high-stakes testing as the exclusive determinant of the accountability of elementary schools in Georgia and the nation should be considered with an understanding of current accepted knowledge about learning capabilities, meaningful test scores, and school improvement prior to the implementation of a new law.

Recommendations

Based upon the findings of this study, several recommendations are made by this researcher. These recommendations include both recommendations for further research,

policymakers, and for educators who are responsible for improved student achievement for all students as measured by high-stakes testing.

1. Policymakers should consider current accepted knowledge and include the insights of Georgia principals before developing legislation for educational reform.
2. Policymakers and local boards of education should consider expansions of the leadership roles of principals as the demand for expertise in high-stakes testing grows, especially in view of the fact that the average years of experience for the principals surveyed was two years.
3. A similar study should be conducted with principals categorized as middle and secondary school principals with a comparison of results across the categories.
4. Further research should be conducted to assess the current accepted knowledge about learning capabilities, meaningful test scores, and school improvement in contrast with the current expectation that every child – regardless of income, gender, race, ethnicity, or disability – can learn with no child being left behind, regardless of the individual starting points.
5. Provide sufficient support for training and consideration of time away from current responsibilities when providing comprehensive change for school improvement.
6. Consider the negative implications of labeling schools. Instead, provide support, structure, and training and the expectation that every school can succeed with no school being left behind.

7. College and university officials should be included in educational reform efforts in view of their responsibility to design degree programs for future principals, teachers, and the challenge to implement effective and efficient school reform.

REFERENCES

- American Educational Research Association (2000). AERA position statement concerning high-stakes testing in prek - 12 education. The Education & Research Network, 1-5. Available online at <http://www.aeara.net/about/policy/stakes.htm>. (January 3, 2006).
- Babbie, E. (2001). The practice of social research (9th ed). Belmont, CA: Wadsworth/ Thomson Learning.
- Benveniste, G. (1985). The design of school accountability systems. Educational Evaluation and Policy Analysis, 7(3), 261-279.
- Bernhardt, V. L. (2001, June 26). Data analysis. Paper presented at the Alpine School District Administrative U-PASS Training, Lindon, Utah.
- Bracey, G. (2000). High stakes testing.
- Broder, D. S. (2001). Long road to reform: Negotiators forge education legislation. The Washington Post, p.A1.
- Brown, D. C. (1993). America 2000 and policy implications for the country. Measurement & Evaluation in Counseling & Development, 26(1), 48-53.
Retrieved February 20, 2004 from PsycINFO database.
- Clancy, K. J. (2000). Making more sense of MCAS scores. Boston Globe, 257(15), A19.
- Cousins, J. B. & Leithwood, K. A. (1992). Enhancing knowledge utilization as a strategy for school improvement. Knowledge: Creation, Diffusion, and Utilization, 14, 305-332.
- Creswell, J. W. (2003). Research design: Qualitative, quantitative, and mixed methods approaches (2nd ed.). Thousand Oaks, CA: Sage.

- Creswell, J. W. (2005). Educational research: planning, conducting, and evaluating quantitative and qualitative research (2nd ed.). Columbus, Ohio: Pearson.
- Denig, S. J., & Quinn, T. (2001). Ethical dilemmas for school administrators. High School Journal, 84(4), 43-49. Retrieved February 6, 2004 from PsycINFO database.
- Dillman, D. A. (1978). Mail and telephone surveys: The total design method. New York: John Wiley.
- Education Commission of the States (ECS). (2001). A closer look: State policy trends in three key areas of the Bush education plan - testing, accountability and school choice (Report). Denver: Education Commission of the States (ECS).
- English, F. W. (2003). The postmodern challenge to the theory and practice of educational administration. Springfield, Illinois: Charles C. Thomas.
- English, F. W., & Hill, J.C. (1994). Total quality education: Transforming schools into learning places. Thousand Oaks, California: Corwin.
- Finn, C. E. (1991). We must take charge: Our schools and our future. New York: The Free Press.
- Finn, C. E., Petrilli, M. J. (2000). The state of state standards 2000 (Report): The Thomas B. Fordham Foundation.
- Finn, C. E., Petrilli, M. J., & Vanourek, G. (1998a). The state of state standards (Report). Washington: Thomas B. Fordham Foundation.
- Finn, C. E., & Walberg, H. J. (1994) Radical education reforms. Berkeley, CA: McCutchan.

- Firestone, W. A. (2004). Educational leaders and assessment-based reform. In S. Mathison & W. E. Ross (Eds.), *Defending Public School* (Vol. 4, pp.137-147). Westport, CT: Praeger Perspectives.
- Fowler, E. J. (2002). Survey research methods (3rd ed.). Thousand Oaks, CA: Sage.
- Fullan, M. (2005). Leadership & sustainability. Thousand Oaks, California: Corwin Press.
- Fullan, M. (2003). The moral imperative of school leadership. Thousand Oaks, California: Corwin Press.
- Fullan, M. (1993). Change forces: Probing the depths of educational reform. New York: Falmer Press.
- Gall, M. D., Borg, W. R., & Gall, J.P. (1996). Educational research: An introduction (6th ed.). White Plains, NY: Longman.
- Gallup, A. M., & Rose, L. C. (2002). The 34th annual Phi Delta Kappa/Gallup poll of the public's attitudes toward the public school. Phi D Georgia Department of Education. (2005). School address download. Retrieved October 15, 2006 from http://app.doe.k12.ga.us/ows-bin/owa/main_pack_school_addr.entry_form.xls
- Georgia Department of Education. (2003). No child left behind act of 2001 (NCLB) Retrieved February 23, 2004, from <http://www/doe.k12.ga.us/support/plan/nclb.asp>
- Glickman, C. D. (1990). Open accountability for the '90s: Between the pillars. *Educational Leadership* (April), 38-42.
- Gravetter, F. J., & Wallnau, L. B. (2000). Statistics for the behavioral sciences (5th ed.).

- Belmont, CA: Wadsworth/Thomson Learning.
- Hadderman, M. (2000). Standards: The policy environment (ERIC Digest EDO-EA-00-04). Eugene: Office of Educational Research and Improvement (ED), Washington, DC.
- Harris, L. (1989). The metropolitan life survey of the American teacher 1989: preparing schools for the 1990's: New York:
- Heubert, J., & Hauser, R. (1999). High stakes testing for tracking, promotion, and graduation. Washington, D.C.: National Academy Press.
- Hochschild, J., & Scott, B. (1998). Trends: governance and reform of public education in the United States. Public Opinion Quarterly 62(1):79-120.
- Hombo, C. M. (2003). NAEP and no child left behind: Technical challenges and practical solutions. Theory into Practice, 42(1),59-65.
- Johnson, J., & Immerwahr, J. (1994). What americans expect from the public schools. New York: Public Agenda.
- Jones, L. V. (2001). Assessing achievement versus high-stakes testing: A crucial contrast. Educational Assessment, 7(1), 21-8. Retrieved February 6, 2004, from Academic Search Premier database.
- Kerlinger, F. N. (1986). Foundations of behavioral research (3rd ed.). Fort Worth, TX: Holt, Rinehart and Winston, Inc.
- Krejcie, R. V., & Morgan, D. W. (1970). Determining sample size for research activities. Educational and Psychological Measurement, 30, 607-610.
- Kohn, A. (2000). The case against standardized testing: Raising the scores, ruining the schools. Portsmouth, NH: Heinemann.

- Kohn, A. (1993). Punished by rewards: The trouble with gold stars, incentive plans, A's, praise and other bribes. Boston: Houghton Mifflin.
- Leithwood, K. A., Seashore, L., Anderson, S., & Wahlstrom, K. (2004). How leadership influences student learning. University of Minnesota, Center for Applied Research and Educational Improvement; University of Toronto, Ontario Institute for Studies in Education and The Wallace Foundation.
- Leman, N. (2000). Don't let scandal scuttle standards. Newsweek, 135, 54.
- Linn, R. L. (2000). Assessments and accountability. Educational Researcher, 29(2), 4-16.
- Lyman, H. B. (1998). Test scores and what they mean (6th ed.). Boston: Allyn & Bacon.
- Mathis, W. (2003). Costs and benefits. Phi Delta Kappan, 84(9), 670-686. <http://www.doe.k12>.
- Miller, D. C., & Salkind, N. J. (2002). Handbook of research design & social
- Ohanian, S. (1999). One size fits few: The folly of educational standards. Portsmouth, NH: Heinemann.
- Pedulla J. J., Abrams, L. M., Madaus, G. F., Russell, M. K., Ramox, M. A. & Miao, J. (2003). Perceived effects of state-mandated testing programs on teaching and learning: findings from a national survey of teachers. Retrieved February 18, 2005 from <http://www.bc.edu/nbetpp/statesments/nrb2.pdf>
- Peterson P. E., & West, M. R. (2003). No child left behind? The politics and practice of school accountability. Washington, D. C.: Brookings Institution Press.
- Popham, J. W. (2001). The truth about testing. Alexandria, Virginia: Association for Supervision and Curriculum Development.

- Popham, J. W., & Lindheim, E. (1981). Implications of a landmark ruling on Florida's minimum competency test. Phi Delta Kappan, 63(1),18-20.
- Porter, A. C. (1998). The effects of upgrading policies on high school mathematics and science. In Brookings Paper on Education Policy, D. Ravitch, ed. Washington, DC: Brookings Institution Press.
- Reeves, D. B. (2000). Standards are not enough: Essential transformations for school success. NASSP Bulletin, 84(620), 5-19.
- Salant, P., & Dillman, D. A. (1994). How to conduct your own survey. New York: John Wiley.
- Schwartz, R. & Gandal, M. (2000). Higher standards, Stronger tests. Education Week, 40.
- Sclafani, S. (2002). No child left behind. Issues in Science and Technology, 19(2), 43-47.
- SPSS (Version 12.0) [Computer Software]. (2003). Chicago: SPSS Inc.
- Stake, R. (1998). Some comments on assessment in U. S. education. Education Policy Analysis Archives (on-line serial) (6:14).
Available:<http://epaa.asu.edu/epaa/v6n14.html>.
- Sutherland, G. (1973). Policy-making in elementary education 1870 - 1895. London: Oxford University.
- The National Commission on Excellence in Education. (1983). A nation at risk: The imperative for educational reform (Report to the Nation and the Secretary of Education, United States Department of Education). Washington, D. C.: The National Commission on Excellence in Education.
- Tunks, J. U. (2001). The effect of training in test item writing on test performance of

- junior high students. Educational Studies, 27(2), 129-142. Retrieved February 6, 2004 from PsycINFO database.
- U. S. Department of Education (2001). No child left behind. Washington, D.C.: Federal Government of the United States. Retrieved from <http://www.ed.gov/nclb/landing.jhtml#>.
- Utah Foundation. (1999). Accountability in public education: an overview and analysis (Report 629). Salt Lake City: Utah Foundation.
- Wallace, D. (1954). A case for—and against—mail questionnaires. Public Opinion Quarterly, 18, 40-52.
- Waters, J. T., Marzano, R. J., & McNulty, B. A. (2003). Balanced leadership: What 30 years of research tells us about the effect of leadership on student achievement. Aurora, CO: Mid-Continent Research for Education and Learning.
- Webb, L. D., Metha, A., & Jordan, K. F. (1992). Foundations of American Education. New York, NY: Macmillan Publishing Company.
- WestEd. (2000). The high stakes of high-stakes testing. Policy Brief. Retrieved January 16, 2005 from <http://www.wested.org>
- Worthen, B. R., Sanders, J. R., & Fitzpatrick, J. L. (1997). Program evaluation: Alternative approaches and practical guidelines (2nd ed.) New York, NY: Addison Wesley Longman, Inc.
- Yudof, M. G., Kirp, D. L., & Levin, B. (1992). Educational policy and the law (3rd ed.). New York, NY: West Publishing Company.

APPENDICES

APPENDIX A
GEORGIA ELEMENTARY PRINCIPALS' PERCEPTIONS OF
HIGH-STAKES TESTING SURVEY

Participant Demographics

1. Years of Experience as a Principal:
 1-5 6-10 11-15 16-20 21-25 26+
2. AYP Status for 2005-06:
 Met Did Not Meet
3. Age:
 <35 35-45 46-55 56+
4. Gender:
 Male Female
5. Educational level:
 M.Ed Ed.S. Ed.D./Ph.D.
6. School Community Served:
 Urban Suburban Rural

The purpose of this survey is to determine elementary principals' perceptions of high-stakes testing. Your responses will help Georgia elementary principals as they seek to improve student achievement. Thank you for your participation in this study.

Please circle the degree to which you agree or disagree with each statement. A one (1) represents Strong Disagreement (SD), while a four (4) represents Strong Agreement.

	Strongly Disagree		Strongly Agree	
1. I should be held accountable for student achievement as measured by high-stakes testing.	1	2	3	4
2. Student achievement will improve as my school receives rewards based on high-stakes testing.	1	2	3	4
3. Student achievement will improve as my school receives sanctions based on high-stakes testing.	4		1	2 3
4. I believe high-stakes testing is fair.	1	2	3	4
5. I believe high-stakes testing improves education for all students.	1	2	3	4

	Strongly Disagree		Strongly Agree	
6. The high-stakes testing results of my school are an appropriate measure of my accountability.	1	2	3	4
7. I consistently use high-stakes testing results to improve student achievement.	1	2	3	4
8. Factors beyond my control influence student achievement and should be considered with my results.	1	2	3	4
9. The purpose of high-stakes testing is to make me a more focused instructional leader.	1	2	3	4
10. My school's high-stakes testing results accurately evaluate my school leadership.	1	2	3	4
11. The test results of my school determine my perception of high-stakes testing.	1	2	3	4
12. The purpose of high-stakes testing is to improve student achievement in my school.	1	2	3	4
13. The purpose of high-stakes testing is to hold principals accountable for their performance.	1	2	3	4
14. How will high-stakes testing influence the way you hold yourself accountable for student achievement in your school?				
15. How has high-stakes testing impacted student achievement in your school?				

APPENDIX B

STUDY PARTICIPANT INFORMED CONSENT LETTER

November 21, 2006

Dear Principal:

I am a practicing Assistant Principal in Toombs County, Georgia, and a doctoral student at Georgia Southern University. I would like to conduct a survey for my dissertation regarding Georgia elementary principals' perceptions of high-stakes testing.

Your response is very important in creating a profile of elementary principals as they address implementation of the current comprehensive education reform in Georgia. Your responses are valuable in that principals are the school leaders with the primary responsibility for meeting Adequate Yearly Progress on an annual basis. Additionally, your responses will be reported in an original study exploring Georgia elementary principals' perceptions of high-stakes testing.

There is no penalty for nonparticipation. You may refuse to respond to any question, or withdraw from the study at any time. Survey completion should require no more than 15 minutes of your time.

Please do not identify yourself on the survey. It will be necessary to document those who responded to the survey for the purpose of contacting nonrespondents. After surveys are received and further contact is established, all coding required for tracking purposes will be destroyed. Coding is used to provide assurance that the researcher is the only one with participant identification. All responses will be treated with utmost confidentiality. Your response indicates that you permit me to use your answers in the study.

You may contact me regarding questions about the study at my office (912) 526-3666, or home (912) 537-8561. If you have questions about your rights as a survey participant in this study, you may contact the Georgia Southern University Office of Research Services and Sponsored Programs at (912) 681-0843.

Thank you in advance for your prompt response. Please use the self-addressed, stamped envelope to return the survey. I will be glad to share survey results with you as requested.

Respectfully,

Sabrina Calhoun
301 S. Broadfoot Blvd.
Vidalia, GA 30474

APPENDIX C

IRB PROPOSAL FORMAT FOR RESEARCH INVOLVING HUMAN SUBJECTS

Research Protocol For Research Utilizing Human Subjects

Personnel:

Sabrina Calhoun, Primary Investigator and Dr. Michael D. Richardson, Advisor will participate in the research.

Purpose:

1. The purpose of this study is to determine the perceptions of Georgia elementary principals about high-stakes testing. The study will address the Georgia elementary principals' perceptions of the demands for high-stakes testing and the appropriateness of the tests as a measure of school effectiveness.
2. The questions to be answered in this experiment include the perceptions of Georgia elementary principals about the demand for high-stakes testing, the purpose of high-stakes testing, if the results account for academic improvement in schools, and if the tests are appropriate measures for accountability. Participants and others will benefit from the knowledge gained in this project because accepted knowledge about learning capabilities, meaningful test scores, and school improvement will be supported. Education reform will benefit from an understanding of the perceptions of elementary principals about high-stakes testing.
3. The current literature reviewed includes the U. S. Department of Education, Georgia Department of Education, Lemann, Linn, Webb, Ydof, Kirp & Levin, Worthen, Sanders & Fitzpatrick, Popham, Finn, and Peterson & West and others. The review included current literature about standardized testing, appropriate use and interpretation of test scores, school reform, accountability of school reform efforts, standards, and the leadership role of principals in school improvement efforts. The

current literature helps frame the hypothesis and research by addressing the history of educational reform efforts within the political and educational context of our country's development, and the positive as well as negative impact of these efforts on student achievement.

Describe your subjects:

A descriptive, survey approach will be employed to address the research questions.

The population for the research is the current 180 school systems existing in Georgia and the 1267 principals listed as heads of elementary schools containing grade groupings incorporation PK – 5 in various combinations. The subjects will be comprised of a random sample with replacement of the 1267 Georgia elementary principals currently listed as the Georgia Department of Education website. A total of 302 elementary school principals will be selected for participation in the study. The approximate ages are 25 – 60. There are no gender requirements or any other identification process that might jeopardize confidentiality. There are no inducements to be used to recruit subjects. A single-stage sampling procedure was used since principal's name and school addresses were listed in the directory.

Methodology (Procedures):

The researcher designed a survey questionnaire to collect data in the investigation to address the research questions. The survey instrument includes 6 demographics questions addressing the years of experience as principal, AYP status, age, gender, educational level, and school community served (urban, suburban, and rural). The survey included 13 items that reflect components of high-stakes testing that are supported in the literature and that address unique facets of NCLB. The questions are

designed to understand principals' perceptions as they implement the high-stakes testing components of NCLB and to address whether elementary principals perceive that specific components of the law are helping their school improve student achievement. In addition, the questions are designed to ascertain whether principals perceive high-stakes testing as a valuable measure of school effectiveness. The study will employ a descriptive, survey approach to address the research questions. A pilot study will be conducted to assess how well the survey addresses issues examined in the study.

Procedures for obtaining informed consent in the methodology:

The informed consent document will be mailed with the survey instrument, Georgia Elementary Principals' Perceptions of High-Stakes Testing Survey, to 608 principals selected for participation in the study. The principals will receive the informed consent document, and by returning the completed survey instrument, will provide documentation of receipt of agreement to the informed consent for this study.

Research involving minors:

This research does not include minors.

Description of possible risk to human subjects:

Study components exclude possible risk to human subjects.

APPENDIX D

IRB REVIEW BOARD APPROVAL FORM

Dear Applicant,

Enclosed below is a copy of your IRB approval letter. You will be receiving a hard copy for your records shortly. If you have any questions or concerns please do not hesitate to contact me.

Sincerely,
Julie B. Cole

=====

To: Sabrina V. Calhoun
301 S. Broadfoot Blvd.
Vidalia, GA 30474

Cc: Dr. Michael D. Richardson, Faculty Advisor
Leadership, Technology and Human Development

From: Julie B. Cole
Office of Research Services and Sponsored Programs
Administrative Support Office for Research Oversight Committees
(IACUC/IBC/IRB)

Date: April 20, 2006

Subject: Status of Application for Approval to Utilize Human Subjects in Research

On behalf of Julie B. Cole, Director of Research Services and Sponsored Programs (IRB), I am writing this letter to inform you that we have completed the review of your Application for Approval to Utilize Human Subjects in your proposed research, "An Analysis of Georgia Elementary Principals' Perceptions of High-Stakes Testing." proposal research project numbered H06153 and it appears that (1) the research subjects are at minimal risk, (2) appropriate safeguards are planned, and (3) the research activities involve only procedures which are allowable. Therefore, as authorized in the Federal Policy for the Protection of Human Subjects, I am please to notify you that the Institutional Review Board has approved your proposed research. This IRB approval is In effect for one year from the date of this letter. If at the end of that time, there have been no changes to the research protocol, you may request an extension of the approval period for an additional year. In the interim, please provide the IRB with any information concerning any significant adverse event, whether or not it is believed to be related to the study, within five working days of the event. In addition, if a change or modification of the approved methodology becomes necessary, you must notify the IRB Coordinator prior to initiating any such changes or modifications. At that time, an amended application for IRB approval may be submitted. Upon completion of your data collection, you are required to complete a Research Study Termination form to notify the IRB Coordinator, so your file may be closed.

Director of Research Services and Sponsored Programs
Georgia Southern University
P. O. Box 8005
Statesboro, GA 30460
P: 912-681-5463 F: 912-681-0719
<http://www2.GeorgiaSouthern.edu/research/Resources/oversight@GeorgiaSouthern.edu>

APPENDIX E
FACE VALIDITY REQUEST

May 15, 2006

Dear:

Thank you for helping me to complete the intermediate phase of my dissertation, by assisting with the preliminary survey critique as a member of an expert panel of judges. Your suggestions will be helpful in assisting me to refine questions that may be confusing and make certain that the questions sufficiently cover the content I intend to explore. Your input will be important in determining face validity of the survey.

The research questions I intend to explore include:

1. What do Georgia elementary principals perceive to be the purpose of high-stakes testing?
2. Do Georgia elementary principals perceive that high-stakes testing improve academic performance?
3. Do Georgia elementary principals perceive that high-stakes testing appropriately measure the effectiveness of their performance?
4. What are the perceptions of Georgia elementary principals regarding rewards and sanctions for high-stakes testing?
5. Do demographic variables make any difference in principal's perceptions of high-stakes testing?

I look forward to your comments. I respectfully request that you return the survey with your remarks included within the next two weeks. This deadline is necessary in order that I may make needed corrections and contact you again if needed, prior to conducting the pilot survey.

Thank you for your time and efforts to contribute to the success of my study. I hope that this effort will contribute valuable information to a variety of audiences regarding the use of high-stakes testing and Georgia elementary principals.

Respectfully,

Sabrina Calhoun
301 S. Broadfoot Blvd.
Vidalia, Georgia 30474

APPENDIX F

PILOT STUDY INFORMED CONSENT LETTER

June 15, 2006

Dear Principal:

I am an elementary assistant principal in Toombs County, Georgia, and a doctoral student at Georgia Southern University. I would like to conduct a pilot survey for my dissertation regarding Georgia elementary principals' perceptions high-stakes testing. The data will provide information regarding elementary principals' perceptions of high-stakes testing.

Your response is very important in this preliminary step of the study. Your response will assist me in ascertaining content validity and reliability of the instrument prior to formally conducting the study. Survey completion should require no more than 15 minutes of your time.

Of course there is no penalty for nonparticipation. You may refuse to respond to any question, or withdraw from the study at any time.

I request that you examine each question for clarity. Please critique each item in regard to its format, language, and content. Additionally, please write what you believe each question means on the survey, so that it can be determined whether each concept is understood as it was intended. I encourage you to provide suggestions and criticisms that will improve the study.

Please do not identify yourself on the survey. It will be necessary to document those who responded to the survey for the purpose of contacting nonrespondents. Coding is used to provide assurance that the researcher is the only one with participant identification. I assure you that all responses will be treated with utmost confidentiality. Your response indicates that you permit me to use your answers in the study.

You may contact me regarding questions about the study at my office (912) 526-3666, or home (912) 537-8561. If you have questions about your rights as a survey participant in this study, you may contact the Georgia Southern University IRB Coordinator at the Office of Research Services and Sponsored Programs at (912) 681-5465.

Thank you in advance for your prompt response. Please use the self-addressed stamped envelope to return the survey. I will be glad to share survey results with you upon request.

Respectfully,

Sabrina Calhoun
301 S. Broadfoot Blvd.
Vidalia, GA 30474

APPENDIX G

STUDY PARTICIPANT INFORMED CONSENT LETTER

December 28, 2006

Dear Principal:

HELP! I recently sent you a survey for a study I am conducting on elementary principals' perceptions of high-stakes testing. I desperately need your response in order that my study may be completed. I am a practicing elementary assistant principal in Toombs County, Georgia and a doctoral student at Georgia Southern University.

Your response is very important and valuable in creating a profile of the perceptions of elementary principals as they implement the high-stakes testing components of NCLB. Your responses will be reported in an original study exploring the perceptions of elementary principals in Georgia.

Of course there is not penalty for nonparticipation. You may refuse to respond to any question, or withdraw from the study at any time. The completion of the survey should not require more than 15 minutes of your time.

Please do not identify yourself on the survey. It will be necessary to document those who responded to the survey for the purpose of contacting nonrespondents. After surveys are received and further contact is established, all coding required for tracking purposes will be destroyed. Coding is used to provide assurance that the researcher is the only person with participant identification. I assure you that all responses will be treated with the utmost confidentiality. Your response indicates that you permit me to use your answers in the study.

You may contact me regarding questions about the study at my office (912) 526-3666, or at my home (912) 537-8561. If you have questions about your rights as a survey participant in this study, you may contact the Georgia Southern University Office of Research Services and Sponsored Programs at (912) 681-0843.

Thank you in advance for providing the assistance that will enable the completion of this study. Please use the self-addressed, stamped envelope to return the survey. I will be glad to share survey results with you upon request.

Sincerely,

Sabrina Calhoun
301 S. Broadfoot Blvd.
Vidalia, GA 30474