



12-2006

Secondary School Principals' Perceptions of the Value and Impact of High-Stakes Tests on Content and Mode of Instruction

Clifford Davis Jr.

University of Tennessee, Knoxville

Recommended Citation

Davis, Clifford Jr., "Secondary School Principals' Perceptions of the Value and Impact of High-Stakes Tests on Content and Mode of Instruction." PhD diss., University of Tennessee, 2006.
https://trace.tennessee.edu/utk_graddiss/4259

This Dissertation is brought to you for free and open access by the Graduate School at Trace: Tennessee Research and Creative Exchange. It has been accepted for inclusion in Doctoral Dissertations by an authorized administrator of Trace: Tennessee Research and Creative Exchange. For more information, please contact trace@utk.edu.

To the Graduate Council:

I am submitting herewith a dissertation written by Clifford Davis Jr. entitled "Secondary School Principals' Perceptions of the Value and Impact of High-Stakes Tests on Content and Mode of Instruction." I have examined the final electronic copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Education, with a major in Educational Administration.

E. Grady Bogue, Major Professor

We have read this dissertation and recommend its acceptance:

Vincent Anfara, Pat Freeland, Glennon Rowell

Accepted for the Council:

Carolyn R. Hodges

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

To the Graduate Council:

I am submitting herewith a dissertation written by Clifford Davis, Jr. entitled "Secondary School Principals' Perceptions of the Value and Impact of High-Stakes Tests on Content and Mode of Instruction." I have examined the final paper copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements of the degree of Doctor of Education, with a major in Educational Administration and Policy Studies.



E. Grady Bogue, Major Professor

We have read this dissertation and recommend its acceptance.



Vincent Anfara



Pat Freeland



Glennon Rowell

Accepted for the Council:



Vice Chancellor and
Dean of Graduate Studies

Thesis
2006b
.Dat

SECONDARY SCHOOL PRINCIPALS' PERCEPTIONS OF THE VALUE AND
IMPACT OF HIGH-STAKES TESTS ON CONTENT AND MODE OF
INSTRUCTION

A Dissertation
Presented for the
Doctor of Education
Degree
The University of Tennessee, Knoxville

Clifford Davis, Jr.

December 2006

Copyright © 2006 by Clifford Davis, Jr.

All rights reserved.

DEDICATION

This dissertation is dedicated to my wife, Angela, and my children, Kevin, Ryan, and Summer, for believing in me, encouraging me, and supporting me in my efforts to reach my personal and professional goals. There were times when I wanted to give up; however, my wife and children would not let me. I am so glad they did not.

ACKNOWLEDGEMENTS

I wish to thank all of those individuals who helped me complete my doctorate in educational administration and policy studies. I would like to thank Dr. E. Grady Bogue, who chaired my committee. I do not believe I would have completed the degree program without his support. I would like to thank Dr. Pat Freeland, Dr. Vincent Anfara, and Dr. Glennon Rowell for serving on my committee. Also, I extend many thanks to Sara Keller and Carol Levine for their assistance in editing my dissertation. Lastly, I would like to thank all of my family and friends for their support, suggestions, and encouragement.

ABSTRACT

Critics of government policies that expand the use of high-stakes tests in public schools claim that these tests have a negative impact on student learning. At the building level, these policies have resulted in a great deal of pressure for educators to raise student performance on these tests.

The purpose of this study was to explore secondary school principals' perceptions of the value and impact of state-mandated tests on content and mode of instruction. The entire population of 541 middle and high school principals from public schools in Tennessee was selected to participate in this study.

Secondary school principals reported agreement on the following issues: (a) high-stakes tests are not an accurate measure of what ESL students know and can do, (b) media coverage of the results of high-stakes tests is unfair to teachers, (c) high-stakes tests are worth the investment of time and money, (d) high-stakes tests have brought attention to education issues, and (e) score differences on high-stakes tests from year to year reflect changes in characteristics in students and not school effectiveness. The principals involved in this study disagreed with the following statements: (a) media coverage of the results of high-stakes tests adequately reflects the quality of education, (b) high-stakes tests motivate unmotivated students, and (c) media coverage adequately reflects the complexity of teaching.

Secondary school principals reported an increase in the amount of time spent on subjects that are part of the state-mandated testing program. In contrast, principals reported a decrease in the amount of time spent on nontested subjects and classroom and student activities.

Significant differences were found in principals' responses when examined by school level. Results indicated that high school principals agreed more than middle school principals did that (a) high-stakes tests motivate previously unmotivated students, (b) high-stakes tests are a fad, and (c) high-stakes tests are not an accurate measure of what ESL students know and can do. Additionally, high school principals across all categories (i.e., urban, suburban, and rural) indicated that their schools spent more time on areas not covered on the state-mandated tests, while middle school principals did not indicate this.

EXTENDED ABSTRACT

The lack of discussion among policymakers and those who implement the policies, as well as the expanded use and consequential nature of state-mandated tests, has fueled arguments by researchers, educators, and others to reverse the growing trend of using high-stakes tests to evaluate the effectiveness of districts, schools, principals, and teachers. The critics of the expanded use of high-stakes tests claim that these tests have a negative impact on teaching and learning in public schools. At the building and classroom levels, these policies have resulted in a great deal of pressure for principals and teachers to raise student performance on these state-mandated tests.

As instructional leaders of schools, principals have the overall responsibility of making decisions and developing policies, procedures, and strategies to address the needs of students and schools under their charge. Therefore, it is important to understand public secondary school principals' perceptions of the value and impact of state-mandated high-stakes tests on content and mode of instruction.

The purpose of this study was to explore those perceptions. Analysis of the study data provided a greater understanding of the impact of high-stakes tests on instruction.

The following were the specific research questions for this study:

1. What are secondary school principals' perceptions of the value of state-mandated high-stakes tests?

2. What are secondary school principals' perceptions of the impact of state-mandated high-stakes tests on the content and mode of instruction?
3. Do secondary school principals' perceptions of the value and impact of state-mandated high-stakes tests on content and mode of instruction differ when examined by school level (e.g., middle and high school) and category (e.g., urban, suburban, and rural)?

The entire population of 541 middle and high school principals from public schools in Tennessee was selected to participate in this study. This number included 272 middle school principals and 269 high school principals. Of the 541 surveys distributed, 310 were returned, for a response rate of 58%.

Secondary school principals reported agreement on the following issues: (a) state-mandated high-stakes tests are not an accurate measure of what ESL students know and can do, (b) media coverage of the results of state-mandated high-stakes tests is unfair to teachers, (c) state-mandated high-stakes tests are worth the investment of time and money, (d) high-stakes tests have brought attention to education issues, and (e) score differences on state-mandated high-stakes tests from year to year reflect changes in characteristics in students and not school effectiveness. The principals involved in this study disagreed with the following statements: (a) media coverage of the results of state-mandated high-stakes tests adequately reflects the quality of education, (b) state-mandated high-stakes tests motivate unmotivated students, and (c) media coverage adequately reflects the complexity of teaching.

Secondary school principals reported an increase in the amount of time spent on subjects that are part of the state-mandated testing program. In contrast, principals reported a decrease in the amount of time spent on nontested subjects and classroom and student activities. Overall, principals' responses indicated that there had been a slight decrease in the amount of time spent in foreign language courses.

Significant differences were found in principals' responses when examined by school level. ANOVAs were run to explore the specific differences in middle and high school principals' responses. Results indicated that high school principals agreed more than middle school principals did that (a) state-mandated tests motivate previously unmotivated students, (b) state-mandated high-stakes tests are a fad, and (c) state-mandated high-stakes tests are not an accurate measure of what ESL students know and can do. Additionally, high school principals across all categories (i.e., urban, suburban, and rural) indicated that their schools spent more time on areas not covered on the state-mandated tests, while middle school principals did not indicate this. Results also revealed that high schools spent more time on foreign language than middle schools did.

TABLE OF CONTENTS

	Page
CHAPTER ONE	1
INTRODUCTION	1
Background	1
Statement of Problem.....	7
Purpose of the Study.....	8
Research Questions.....	9
Definitions.....	9
Assumptions.....	11
Limitations	12
Delimitations.....	13
Significance of the Study	13
Organization of the Study	14
CHAPTER TWO	15
REVIEW OF LITERATURE	15
Introduction.....	15
The Evolution of Accountability	16
Brown versus Board of Education	16
The National Defense Education Act of 1958	17
Elementary and Secondary Education Act	17
The National Assessment of Educational Progress	18
No Child Left Behind Act of 2001	18
Tennessee’s Accountability System	19
Tennessee’s Financial Accountability	21
Tennessee Value-Added Assessment System.....	26
Brief History of High-Stakes Tests.....	27
Nature of Standardized Tests.....	28
Functions of Standardized Tests	29
High-Stakes Testing and Teachers’ Philosophies of Education	29
Teachers’ Response to High-Stakes Testing.....	31
Principals’ Response to High-Stakes Testing	34
Consequences of High-Stakes Testing.....	37
Positive Effects	38
Neutral and Negative Effects	39
Professional Development	46

The Principal as Instructional Leader	48
Summary	51
CHAPTER THREE	52
METHODOLOGY	52
Purpose and Research Questions	52
Research Design	53
Participant Selection	53
Instrumentation	53
Data Collection	60
Data Analysis	61
CHAPTER FOUR.....	63
RESEARCH FINDINGS	63
Introduction.....	63
Analysis of Survey Demographics.....	64
Analysis of Results	72
Research Question 1	72
Research Question 2	74
Research Question 3	77
Summary	92
CHAPTER FIVE	95
SUMMARY, FINDINGS, CONCLUSIONS, DISCUSSION, AND RECOMMENDATIONS	95
Summary	95
Purpose and Research Questions	97
Summary of Findings.....	98
Conclusions.....	102
Discussion	103
Recommendations for Future Research and Practice	106
REFERENCES.....	109
APPENDICES.....	126
APPENDIX A.....	127
Survey Modification Form.....	128
APPENDIX B.....	129
Prenotification Letter	130
APPENDIX C.....	131
Sample Letter to School Principals.....	132
APPENDIX D.....	133
Permission to Use Survey Instrument.....	134
APPENDIX E.....	135
Survey Instrument	136

APPENDIX F	140
Characteristics of Survey Respondents.....	141
APPENDIX G.....	142
Thank You/Reminder Postcard.....	143
APPENDIX H.....	144
Second Sample Letter to Principals	145
APPENDIX I.....	146
Content of Instruction	147
APPENDIX J.....	148
Mode of Instruction.....	149
VITA.....	150

LIST OF TABLES

	Page
Table 1. Tested Areas, Noncore Content, Classroom Activities Scales	57
Table 2. Scale Reliability	60
Table 3. Category.....	65
Table 4. School Performance.....	66
Table 5. School Level	66
Table 6. School Location	67
Table 7. School Size	68
Table 8. Participants' Years of Administrative Experience	69
Table 9. Participants' Gender	69
Table 10. Participants' Age	70
Table 11. Participants' Race.....	71
Table 12. Perceived Value of State-Mandated Test Scale.....	73
Table 13. Tested Areas Scale.....	76
Table 14. Noncore Areas Scale.....	76
Table 15. Activities Scale	78
Table 16. MANOVA Results by Level and Category on the Perceived Value Scale	79
Table 17. ANOVA for Perceived Value of State-Mandated Test Scale Tests of Between- Subjects Effects.....	80
Table 18. Perceived Value Scale Means by Level.....	81

Table 19. Means for Differences in Perceived Value Scale	83
Table 20. MANOVA Results by Level and Category on Tested Areas Scale	83
Table 21. ANOVA for Tested Areas Scale by Level.....	84
Table 22. Means for Differences in Tested Areas Scale.....	84
Table 23. MANOVA Results by Level and Category on Noncore Scale.....	85
Table 24. ANOVA for Noncore Areas Scale Tests of Between-Subjects Effects.....	86
Table 25. Means for Differences in Noncore Areas Scale.....	87
Table 26. Means for Differences in Foreign Language by Category and Level.....	88
Table 27. MANOVA Results by Level and Category on Activities Scale.....	90
Table 28. ANOVA for Activities Scale by Level Tests of Between-Subjects Effects	91
Table 29. Activities Scale Means by Level	91

CHAPTER ONE

Introduction

Background

Some scholars (Bracey, 2003b; Rentner & Hamilton, 2003) have predicted that many of our nation's public schools may find themselves labeled as schools in need of improvement by the year 2014 because of annual test scores. Under current federal and state accountability policies, all students, including special education and English as a second language (ESL) students, must be proficient on state-mandated tests in reading and mathematics. Schools that fail to meet the performance targets risk sanctions and possible closure (Bracey; Rentner & Hamilton).

Accountability has become one of the most debated policy issues in public education today. The current emphasis on accountability in American education is an outgrowth of a number of reform movements in public education (Ahearn, 2000). These reforms have flowed from a growth in public perception, right or wrong, that public schools are performing poorly or are inadequate to meet the demands of an increasingly global society. Over the last 30 years, these reforms have led to a movement of enormous magnitude. *A Nation at Risk*, published in 1983, did a great deal to fuel the movement for accountability in America's public schools. The National Commission on Excellence in Education noted in this report that there were some serious deficiencies in public

schools, which were possibly endangering not only the future of our youth, but also our existence as a nation. The authors of *A Nation at Risk* wrote,

If an unfriendly foreign power had attempted to impose on America the mediocre educational performance that exists today, we might well have viewed it as an act of war. As it stands, we have allowed this to happen to ourselves. (National Commission on Excellence in Education, 1983, p. 5)

A Nation at Risk was one of the first alarms for reform in America's public schools; however, some researchers, writers, and educators believe that this was a false alarm. For example, Bracey (2003a) wrote, "It has been 20 years, though, since *A Nation at Risk* appeared. It is clear that it was false then and is false now" (p. 621). Additional writers and researchers also disagreed with the findings of *A Nation at Risk* and with other researchers who reported the nation's public schools as failing. For instance, Carson, Huelskamp, and Woodall (1993) conducted a study called the Sandia Report in April 1992. The study was suppressed because its quantitative findings were controversial and completely contradicted the findings of those researchers who reported that public education in the United States was in crisis. The researchers focused on education in New Mexico and examined dropout statistics, standardized tests, postsecondary studies, educational funding, international comparisons, and educator status. The researchers found that on nearly every measure, there was a steady or slightly improving trend. For example, Carson et al. (1993) found that America's on-time graduation rate for the past 20 years had been steady at 75% to 85%. Standardized test

scores on the National Assessment of Educational Progress (NAEP) and the Scholastic Aptitude Test were found to have been steady or improving. Additionally, almost 60% of the students attempted postsecondary studies, and nearly 70% of these enrolled in 4-year institutions. The researchers found that nearly 30% of the youth obtained at least a bachelor's degree. These findings completely contradicted research (Morrison, 1991; Stedman, Irwin, Lyke, & Riddle, 1990) that reported America's public schools as declining.

The call for reform at the middle and high school levels has not gone completely unnoticed. In the early 1970s, the idea that young people between the ages of 10 and 14 needed a special kind of school resulted in the formation of today's middle schools. One of the core tenets of middle level education is that students should be grouped into small, personalized environments or teams to meet both the affective and academic needs of middle school students. Other tenets include team teaching, curriculum integration, advisory programs, and student choice in the everyday school experience. When implemented appropriately, middle level education has much to teach high schools about reform.

At the high school level, teams of educators, parents, and concerned citizens have begun to experiment with a variety of strategies to reform America's high schools. Some groups have tried to improve instruction and make curriculum more relevant to student needs using existing school structures. Other groups have been calling for a complete overhaul of existing high school structures. These groups are trying to devise structures

where oversized high schools feel smaller by dividing them into smaller thematic units, programs, or houses. There are others creating new schools—charter schools, independent schools, and new public and private schools (Daniels, Bizar, & Zemelman, 1998, 2001).

The No Child Left Behind Act (NCLB) of 2001 is a landmark piece of education reform designed to improve both teaching and learning for all students in America's public schools. NCLB has set the goal of having every child reach the *proficient* level (as defined by each state) on state-mandated tests by 2013-2014. To achieve this objective, states have developed benchmarks to measure student progress and to ensure that every child is learning. States are required to disaggregate student achievement data and hold schools accountable for subgroups of students such as African American, Hispanic, low socioeconomic status, and special education. This is done to ensure that no student falls through the cracks (U.S. Department of Education, 2004).

Accountability is the driving force behind NCLB. This legislation requires that all states develop accountability plans to measure both student achievement and the effectiveness of public schools. By now, every state has adopted a statewide accountability system except Iowa, where local districts have developed their own standards and benchmarks. These accountability systems have been written into state statutes. As a result, state departments of education and local school boards have rewritten their regulations and policies to comply with state accountability systems (Patterson, 2002; Rentner & Hamilton, 2003).

A major component of state accountability systems is state-mandated testing. The results of state-mandated tests are used for a variety of policy purposes. These include providing evidence of school effectiveness, assessing the achievement of local and state standards, accrediting school districts, and evaluating teacher effectiveness, to name a few (Brown, 1993; Patterson, 2002). Because state-mandated tests are used to make decisions such as these, they have been labeled as *high-stakes tests*. Educators cannot afford to ignore the issues surrounding state-mandated tests or high-stakes tests because they will be affected by how well their students perform on these tests (Brown).

Tennessee has been one of the leaders in the nation's call for accountability in public education. In 1992, the Tennessee legislature passed the Education Improvement Act (EIA). This legislation increased the amount of money spent on public schools in Tennessee; however, it included an unconventional accountability component to ensure that, for these additional dollars, the students in the state would be provided a quality education. The accountability component in the EIA is called the Tennessee Value-Added Assessment System (TVAAS) and was developed by Dr. William Sanders, while a faculty member at the University of Tennessee. Until the passage of NCLB, it was the heart of Tennessee's accountability system. TVAAS is a statistical method of determining the effectiveness of teachers, schools, and districts on students' learning in Grade 3 through Grade 8 in mathematics, science, social studies, language arts, and reading.

NCLB required that all states establish an accountability plan to hold all schools and school districts accountable for students' academic growth. Even though Tennessee had an accountability system prior to NCLB, the federal government still required that the state submit a new plan that would meet the requirements of No Child Left Behind. Prior to NCLB, only norm-referenced tests were used as a basis for TVAAS. Provisions in NCLB required Tennessee to use both criterion-referenced and norm-referenced components to determine whether students are proficient in reading, language arts, and mathematics for Grades 3, 5, and 8. The state submitted a plan and received official approval in May 2003. Tennessee's accountability plan meets the requirements of NCLB and the state requirements in the EIA (Tennessee Department of Education, 2003a).

The results from the assessments are used by the state to determine the progress of schools and districts toward meeting annual statewide goals in reading, language arts, and mathematics. These components are used to determine the districts' and schools' yearly progress toward achieving Tennessee's academic standards. Adequate yearly progress (AYP) is the minimum level of improvement that states, districts, and schools must achieve each year (Tennessee Department of Education, 2003a). Districts and schools that fail to meet AYP are subject to sanctions and possible closure. Those schools and districts that exceed AYP requirements are rewarded for their efforts.

A reading of the literature seems to indicate that the public's perception of American education has declined over the past 30 years. To combat the public's perception of public schools as failing, policymakers have mandated the use of high-

stakes tests to hold students, schools, educators, and school systems accountable. Depending on the state, scores from these tests may be used to make high-stakes decisions about students, educators, schools, and school districts (e.g., promotion or graduation for students, cash rewards for teachers, accreditation decisions for schools, and operating autonomy or funding for districts). The passage of NCLB in 2001 further increased accountability at the school level.

Statement of Problem

Since the publication of *A Nation at Risk* in 1983, policymakers have come under increased pressure from the public and the federal government to improve student performance at the secondary school level. As a result, federal, state, and local policymakers have enacted policies such as NCLB and the EIA to improve student performance. To achieve this goal, policymakers have pushed for increased use of high-stakes tests, school report cards, and rewards and sanctions for districts and schools to produce higher student performance.

The lack of discussion among policymakers and those who implement the policies, as well as the expanded use and consequential nature of state-mandated tests, has fueled arguments by researchers, educators, and others to reverse the growing trend of using high-stakes tests to evaluate the effectiveness of districts, schools, principals, and teachers. The critics (Bracey, 2003b; Kohn, 2001; Madaus, 1988) claim that these tests have a negative impact on teaching and learning in public schools. At the building

and classroom levels, these policies have resulted in a great deal of pressure for principals and teachers to raise student performance on these state-mandated tests.

Clarke and Gregory (2003) asserted that educators are faced with two choices: They can let noneducators determine the educational agenda, or they can contribute their expertise to the debate that is taking place at the local, state, and national levels. The principal is responsible for encouraging the use of effective instructional methods and discouraging the use of instructional methods that do not improve teaching and student learning. As the instructional leader of the school, the principal has the overall responsibility of making decisions and developing policies, procedures, and strategies to address the needs of students and schools under their charge (Blase & Blase, 2004). Therefore, it is important to understand public secondary school principals' perceptions of the value and impact of high-stakes tests on content and mode of instruction. Moreover, there has been very little research that examines principals' perceptions of the impact of high-stakes tests on instruction.

Purpose of the Study

This study explored secondary school principals' perceptions of the value and impact of state-mandated high-stakes tests on content and mode of instruction in Tennessee.

Research Questions

The study was guided by the following questions:

1. What are secondary school principals' perceptions of the value of state-mandated high-stakes tests?
2. What are secondary school principals' perceptions of the impact of state-mandated high-stakes tests on the content and mode of instruction?
3. Do secondary school principals' perceptions of the value and impact of state-mandated high-stakes tests on content and mode of instruction differ when examined by school level (e.g., middle and high school) and category (e.g., urban, suburban, and rural)?

Definitions

To assist the reader in evaluating the research, the following definitions are provided to ensure uniformity and understanding of these terms throughout the study. These definitions will aid the reader in gaining an explicit understanding of what each term means in the context of the problem.

1. **Accountability**—The process by which schools are held responsible for the academic achievement of students and the job performance of educators. This process includes standardized testing, reporting of results to the public, the use of rewards, and often the imposition of sanctions (Patterson, 2002).

2. Adequate yearly progress (AYP)—An individual state’s measure of yearly progress toward achieving state academic standards. AYP is the minimum level of improvement that states, school districts, and schools must achieve each year (Tennessee Department of Education, 2003b).
3. Content of instruction – Topics and skills for a grade level; The items operationally defining content of instruction are listed in Appendix I.
4. Content validity—The degree to which a test measures an intended content.
5. Criterion-referenced test—A test that is designed to measure students’ knowledge or mastery of various instructional objectives (Bersola, 2002; Mehrens & Lehmann, 1987).
6. High-stakes tests—Tests used for the certification of teachers, promotion of students from one grade to the next, awarding of high school diplomas, assignment of students to remedial classes, allocation of funds to a school district, awarding of merit pay to teachers on the basis of their students’ performance, certification or recertification of a school district, and placement of school systems into “educational receivership” (Madaus, 1988).
7. Mode of instruction – Refers to a variety of instructional methods used by educators; The items operationally defining mode of instruction are listed in Appendix J.

8. Norm-referenced test—A test that compares an individual student’s score to the average performance of a reference group (norm). It ranks students in order of achievement (Bersola, 2002; Mehrens & Lehmann, 1987).
9. Performance level—The level at which a school’s students score on state-mandated tests relative to students at other schools in the state (e.g., below average, average, or above average). This information will be reported by the respondents; however, the researcher will verify reports by checking the state’s Web site containing assessment data.
10. Secondary school—A school that includes any combination of Grade 6 through Grade 12 in the public school system.
11. Standardized test—A commercially prepared test. It provides methods for obtaining samples of behavior under uniform procedures; that is, the same fixed set of questions is administered with the same set of directions and timing constraints, and the scoring procedure is carefully delineated and kept constant. Scoring is usually objective (Mehrens & Lehmann, 1987).
12. State-mandated tests—Standardized tests that a state requires of its schools at specific grade levels (Clarke, Arnold, Rhoades, Abrams, & Li, 2003).

Assumptions

The following are assumptions that were made while designing this survey research:

1. It was assumed that all participants would answer all survey questions honestly and to the best of their abilities.
2. The survey would enable the researcher to collect comparable data from all participants.
3. The participants would complete the survey rather than assigning it to another staff member.
4. The participants would understand that their responses would be kept confidential.

Limitations

Limitations of this study include the following:

1. The use of a survey, while allowing large sampling, does not furnish the opportunity for more in-depth comments.
2. The instrument used to collect data for this study was a survey. The instrument was mailed to the participants in the study. Mailing the survey may have reduced the response rate of the participants, thereby reducing the generalizability of the study.
3. There is no assurance that the participants were truthful in their responses.

Delimitations

Delimitations of the study include the following:

1. The study was confined to public secondary school principals' perceptions of the value and impact of state-mandated high-stakes tests on content and mode of instruction.
2. The study was restricted to secondary school principals in Tennessee.
3. In order to assure the manageability of the collected data, the researcher used only multiple-choice items and did not include open-ended response items.

Significance of the Study

In examining the current research, the researcher found many studies that examined teachers' and principals' perceptions of the impact of state-mandated tests on teaching and learning (Abrams et al., 2003; Adams & Karabenick, 2000; Amrein, 2002; Barksdale-Ladd & Thomas, 2000; Bersola, 2002; Brown, 1993; Cooley & Shen, 2003; Gordon & Reese, 1997; Kaplan & Owings, 2001; Mitchell, 1996; Vogler, 2002; Williams, 2002; Young, 1996). Only two of these studies (Brown; Young) were conducted in the state of Tennessee. None of these studies were done after the May 2003 implementation of Tennessee's new accountability policy as required by NCLB.

The goal of this proposed study is to build on previous research related to the impact of high-stakes tests and provide a statewide picture of how public secondary school principals in Tennessee perceive the value and impact of high-stakes tests on

content and mode of instruction under Tennessee's current accountability policy. Results from this survey will be available to policymakers for their use in considering the nature of the state of Tennessee's testing program.

Organization of the Study

Chapter One has presented the introduction; statement of the problem; purpose of the study; research questions; definition of terms; assumptions, limitations, and delimitations of the study; and significance of the study. Chapter Two contains the review of literature and research related to accountability and high-stakes tests. The methodology and procedures used to gather and analyze data for the study are presented in Chapter Three. The results and analyses and findings from the study will be contained in Chapter Four. Chapter Five will include a summary of the study and a discussion of conclusions and recommendations for further research.

CHAPTER TWO

Review of Literature

Introduction

High-stakes testing has been the driving force behind educational reform efforts in the United States for more than a decade. Policymakers have mandated the use of these tests to gather information about both the quality of teaching and the level of student learning in public schools as a means of holding students, educators, schools, and school systems accountable. In 2001, passage of NCLB moved accountability in public schools to a new level of national visibility. This federal mandate requires states to bring all student groups to the proficiency level in reading and mathematics on state tests by 2014. The mandate also holds states responsible for making adequate yearly progress toward this goal.

Even though state testing programs are now being used in all 50 states, there is still a debate as to whether they actually improve teaching and student learning in the nation's schools. The lack of discussion between policymakers and those who implement the policies, as well as the expanded use and consequential nature of these tests, has led to heated debates by researchers, educators, and others that call for a reexamination of the use of high-stakes tests in evaluating the effectiveness of districts, schools, principals, and teachers.

The focus of this literature review was limited to the impact of high-stakes tests on teaching and learning, due to the broad nature of the subject. Some studies focused on elementary schools, while others focused on high school exit exams. The literature review includes several categories: the evolution of accountability, Tennessee's accountability system, a brief history of high-stakes tests, high-stakes testing and teachers' philosophies of education, the nature of standardized tests, functions of standardized tests, teachers' response to high-stakes testing, principals' response to high-stakes testing, consequences of high-stakes testing, and professional development.

The Evolution of Accountability

In order to understand the accountability movement from a historical perspective, it is necessary to understand the context in which it occurred. The following is a summary of the major events of the movement.

Brown versus Board of Education

In 1954, the Supreme Court ruled on *Brown v. Board of Education* (n.d.). In this case, Oliver Brown, a black railroad worker of Topeka in Shawnee County, Kansas, sued the Board of Education for not allowing his daughter, Linda Brown, to attend an all-white school that was located in their neighborhood. The case, however, was about much more than that. *Brown v. Board of Education* addressed whether the state had the right to sustain separate but equal institutions (including schools) that segregated black Americans into a world with far fewer opportunities than white Americans had available.

The Supreme Court's decision outlawed racially separated institutions and allowed blacks equal access to all government-funded institutions. As a result of *Brown v. Board of Education*, the federal government poured money into public education in an effort to produce greater academic gains for African-American students who had been denied access to a quality education for so many years ("Brown 40 Years On," 1994; *Brown v. Board of Education.*; Graglia, 1996; Halberstam, 2004).

The National Defense Education Act of 1958

In October 1957, the Soviet Union launched the first man-made satellite, Sputnik. This was a shock to Americans because the U.S. had assumed preeminence in science. As a result of Sputnik, America began to question the effectiveness of its educational system. To regain preeminence in science, the federal government passed the National Defense Education Act (NDEA) of 1958. The purpose of this act was to bring American education to a level consistent with the needs of the country economically, militarily, and politically. It was intended to reform public education and create schools able to produce the best and brightest scientists and mathematicians. As a result, it was thought that America would regain its economic, military, and political superiority (Freund, 2002; Sidney, 2003).

Elementary and Secondary Education Act

In 1965, President Johnson signed into law the Elementary and Secondary Education Act (ESEA). This legislation was meant to improve public schools and help poor children climb out of poverty by providing them with a better education. Passage of

ESEA, which drastically expanded the federal government's role in education, produced more federal funds for local school districts. Billions of dollars were spent to implement educational programs for disadvantaged students. These programs were designed to raise disadvantaged students to academic levels that would allow them to compete with more affluent students and to provide them with a quality education. ESEA had a direct influence on the development of state and local testing programs; under the provisions of ESEA, standardized tests were to be used to evaluate the effectiveness of these educational programs and to provide some measure of accountability (Rentner et al., 2003).

The National Assessment of Educational Progress

The National Assessment of Educational Progress (NAEP), also known as "The Nation's Report Card," was established in 1969 with program funding from the U.S. Office of Education and the National Center for Education Statistics. It is the only national assessment that provides cumulative data and identifies trends in the educational achievement of students in public schools. The NAEP assesses achievement of students at 9, 13, and 17 years of age and has been an indicator of what American students know and can do in the areas of mathematics, reading, science, writing, U.S. history, civics, and the arts (Patterson, 2002; U.S. Department of Education, 2003).

No Child Left Behind Act of 2001

In 2001, the federal government made a great deal of progress in its efforts to improve standards and achieve more accountability in public schools when both the

House of Representatives and the Senate passed President George W. Bush's No Child Left Behind (NCLB) Act by very wide margins. This legislation received unprecedented support from both the Democratic and Republican parties. At the heart of this legislation is the use of standardized testing in reading, math, and science to determine how effective school districts, schools, administrators, and teachers are in educating the students in America's public schools. In addition to measuring student progress, NCLB is meant to be a catalyst for change in the culture that currently exists in U.S. public schools (Hill, 2003).

Tennessee's Accountability System

Tennessee has a single, unitary, statewide accountability system for all public schools. Tennessee's plan meets the federal government's requirements under NCLB. This system requires standards-based assessments in reading, language arts, and math for Grades 3, 5, and 8. It has both norm-referenced and criterion-referenced components. Both types of tests are required to determine AYP. The state established its starting points based on spring of 2003 data and set intermediate goals and annual measurable objectives to ensure that all Tennessee students are proficient in reading, language arts, and mathematics by 2013-2014 (Stronge & Tucker, 2000).

Under the state's accountability system, there are two grade spans in the accountability system: elementary/middle (K-8) and high school (9-12). All schools, regardless of grade level, must meet the following criteria to demonstrate AYP:

- Math (95% participation rate on the assessment for all students and each subgroup, and meet the annual measurable objective for math)
- Reading/Language Arts (95% participation rate on the assessment for all students and each subgroup, and meet the annual measurable objective for reading/language arts)
- Additional Indicator (attendance rate for elementary/middle school and graduation rate for high school). (Tennessee Department of Education, 2003b, p. 2)

Tennessee's schools and districts must meet annual measurable objectives in reading/language arts and math for all students, including all of the required subgroup populations (race/ethnicity, economically disadvantaged, disabled students, and limited English proficiency or ESL students) that have 45 or more students. If a school or district fails to meet any one of its annual measurable objectives for all of its students or any one of its subgroups, it may fail to meet AYP. Schools are allowed to demonstrate that they have made AYP by using the most current year, the most current 2-year average, or the most current 3-year average.

Tennessee's system includes both rewards and sanctions, respectively, for those schools and districts that meet or fail to meet the state's requirements. Rewards for schools and districts that meet the state's required annual measurable objectives include recognition and possibly monetary awards. Sanctions range from being identified to

having alternative governance or takeover by the state (Tennessee Department of Education, 2003).

Tennessee's Financial Accountability

In the mid-1980s, the country's confidence in public education began to wane. In part, this was due to the nation losing its competitive edge in the economic market and also the publication of *A Nation at Risk*. This publication warned that the nation was jeopardizing its future because of the lack of standards and accountability in public schools (Sanders & Horn, 1998).

Under the leadership of Governor Lamar Alexander, Tennessee moved to improve education with the passage of the Comprehensive Education Reform Act (CERA) of 1984. CERA included a major increase in educational spending and a career ladder system for teachers. The teacher evaluation system devised for the upper levels of the career ladder was performance based.

As a result of the publication of *A Nation at Risk*, President George H. Bush convened an education summit in 1989 with the governors of the 50 states to address the problems facing education in the United States. This group also had the task of developing a plan of action for addressing these problems (Stronge & Tucker, 2000). As a result, the governors developed six national education goals:

1. All children in America will start school ready to learn.
2. The high school graduation rate will increase to at least 90%.

3. American students will leave Grades 4, 8, and 12 having demonstrated competency in challenging subject matter, including English, mathematics, science, history, and geography, and every school in America will ensure that all students learn to use their minds well, so they may be prepared for responsible citizenship, further learning, and productive employment in our modern economy.
4. U.S. students will be first in the world in mathematics and science achievement.
5. Every adult American will be literate and will possess the knowledge and skills necessary to compete in a global economy and exercise the rights and responsibilities of citizenship.
6. Every school in America will be free of drugs and violence and will offer a disciplined environment that is conducive to learning.

Congress later amended this original list to include two more goals:

7. By the year 2000, the nation's teaching force will have access to programs for the continued development of their professional skills and the opportunity to acquire the knowledge and skills needed to instruct and prepare all American students for the next century.
8. By the year 2000, every school will promote partnerships that will increase parental involvement and participation in promoting the social,

emotional, and academic growth of children. (as cited in DuFour & Eaker, 1998, p. 5)

How these goals were to be accomplished was left to each state. These governors went back to their home states and challenged their legislators to develop legislation that would achieve these national goals for education. The resulting legislation varied from state to state but had some commonalities. It all called for higher academic standards and greater accountability linked to assessment of educational outcomes.

In 1989, discussions were initiated among Tennessee legislators that eventually led to the enactment of the Education Improvement Act (EIA) of 1992 in Tennessee. This piece of legislation occurred because of a lawsuit filed by 66 small school systems against the state. These systems argued that Tennessee's funding system denied some students the right to an adequate free education and violated the equal protection provision of the state's constitution (Ceperley & Reel, 1997; Smith, Detch, & Morgan, 2004).

Historically in Tennessee, small community-based businesses had been pretty evenly dispersed throughout the state. Tennessee has no state income tax and relies on sales tax to support its programs. Since these small businesses were evenly distributed, all school districts received an equitable piece of the pie when it came to funding of their schools. With the advent of large malls and discount superstores, many small community-based businesses were forced to close.

Usually these malls and discount superstores were located in larger metropolitan

areas. This resulted in a shift in tax revenues; the wealthy school districts got wealthier and the poor ones poorer. In some cases, the larger, more affluent school districts were spending twice as much per student as some of the smaller school systems that had fewer resources (Ceperley & Reel, 1997).

Under Tennessee's funding system in 1988, schools in smaller, less affluent districts offered far less to students than schools in wealthier districts. Schools in smaller districts did not have equal access to adequate laboratory facilities, computers, current and new textbooks, adequate buildings, advanced placement courses, varied curriculum, advanced foreign language courses, music and art courses, or drama and television courses. Smaller school districts were not able to retain teachers, fund needed administrators, and provide sufficient physical education and other programs (Ceperley & Reel; Smith et al., 2004).

The wealthier districts offered a wide variety of advanced placement courses; a broad curriculum with advanced science and math courses; adequate labs in both junior high and high schools; a choice of foreign languages; multiple computer courses; art, music, and drama courses; sufficient and current textbooks; and adequately supplied libraries. The schools were newer, cleaner, safer, and provided an environment conducive to learning (Ceperley & Reel, 1997).

As a result of *Tennessee Small School Systems, et al. v. McWhorter* (1993), policymakers wanted to avoid using the "Robin Hood" strategy of redistributing existing resources: taking from the more affluent districts and giving to the poor, rural districts.

The politicians wanted to solve their funding dilemma by raising taxes and gradually raising the level of funding for the poorer districts until equity was reached. The politicians knew that raising taxes would be popular among educators but not among the general public. In order to raise taxes, legislators needed the support of the business community (Ceperley & Reel, 1997; Smith et al., 2004).

In return for their support, business interests, led by the Tennessee Business Roundtable, demanded accountability. First, business wanted accountability for results at the district level. All superintendents would have to be appointed by elected boards instead of by the public. Business argued that school systems needed a single point of accountability.

(Ceperley & Reel, p. 134)

The business community also wanted accountability results at the school-building level and argued that principals should have performance contracts and should be held accountable if they failed to achieve the agreed-on results. Business also wanted accountability at the classroom level, lower dropout rates, higher promotion rates, and increased student achievement (Ceperley & Reel, 1997).

Early drafts of the EIA contained all of the accountability components demanded by the business community except a way of linking student progress to the classroom. According to Ceperley and Reel (1997), when legislators searched to fulfill this requirement, they discovered the value-added assessment model. This model was based on studies done by Dr. William Sanders while at the University of Tennessee.

Tennessee Value-Added Assessment System

The Tennessee Value-Added Assessment System (TVAAS) was the first accountability system of its kind to be used statewide. It applied a new yardstick to district, school, and teacher performance. The system has been praised by some and questioned by others. The heart of the TVAAS is the collection of longitudinal test data on every child in the Tennessee public schools. TVAAS contains approximately 5 million student records. For most students, several years of test data exist to use in estimating normal learning gains from year to year. Students' previous academic growth becomes the standard for future growth. TVAAS is like a pretest/posttest system; the pretest and posttest data are the student's scale score from the previous school year and the current school year, respectively. Determining the amount of learning gains for any given year is a matter of subtracting the previous year's scale score from the current year's scale score. Value-added takes the gain each student makes each year and compares it to the gain made by a normative sample for that subject between the two grade levels. If the normal gain from fifth to sixth grade in science was 12 points, a sixth-grade teacher's students who averaged a 16-point gain for the year would score 133, or 133% of the normal gains. A teacher whose students averaged an 18-point gain would score 150% (Stronge & Tucker, 2000).

Brief History of High-Stakes Tests

As more and more pressure is being placed on public schools and students to produce better results on high-stakes tests, the use of testing is becoming more widespread. During the last century, the uses of high-stakes testing have grown. Formal testing has become an accountability tool in American society.

Standardized testing has been a part of American public schools for over a half century. It initially served as a way of comparing schools and students and was also used to sort students according to their abilities. Test scores were used to identify those capable of succeeding in higher level education and those who were better suited for vocational school (Amrein & Berliner, 2002; Walker, 2000).

The 1970s ushered in the era of minimum competency testing. Reformers looked to improve education and ensure that all students were accountable for achieving basic standards by using tests to determine grade-level promotions and eligibility for high school graduation. By the early 1980s, almost three fourths of the states had minimum competency exams. These exams took the form of multiple-choice tests that students either passed or failed. These exams did very little to measure how much students had learned or how advanced they were (Amrein & Berliner; Clarke, Haney, & Madaus, 2000; Walker, 2000).

The growing criticism of public schools by policymakers and the public in the late 1980s contributed to the increasing use of assessment tied to accountability for student and school performance. During the early days of the accountability movement, test

results were criticized as showing inflated patterns (Walker, 2000). Because of the high stakes tied to these tests, critics such as Kohn (2001) asserted that teachers were teaching to the test.

Today, reformers emphasize accountability as tied to state standards. These standards detail what students should know and how well they should perform. States now align their assessments with these standards. As a result, they are requiring more from the teachers and students in public schools (Walker, 2000).

Nature of Standardized Tests

Historically, American schools have used commercially developed tests to diagnose and classify students. Test scores have been used to place students into different curricular and instructional tracks, to retain them, and to assign them to remedial and special education programs (Madaus, 1991).

Most of these commercially developed tests fall into one of two categories: norm-referenced or criterion-referenced. In norm-referenced tests, an individual student's score is compared to the average performance of a reference group (norm). This type of test ranks students in order of achievement. A criterion-referenced test describes how students perform on tasks that have specified content objectives. Criterion-referenced tests do not compare the performance of individual students to the performance of others (Bersola, 2002; Mehrens & Lehmann, 1987).

Functions of Standardized Tests

Standardized tests have varied functions that can be summed up by saying they assist educators, students, parents, and others in making decisions. These various groups make numerous decisions; the more informed they are, the better their decisions will be (Mehrens & Lehmann, 1987).

Standardized tests are used to allocate rewards and sanctions, to measure student learning, to evaluate the effectiveness of programs, and to place students. For example, Amrein and Berliner (2002) examined the high-stakes testing programs in 18 states to determine how they were affecting student learning. They also analyzed the 45 states that have utilized these tests in accountability programs. Their analysis of the 45 states that hold schools accountable for test scores revealed that all 45 states hold schools accountable by publishing school or district report cards. Among these states, 27 hold schools accountable through rating and ranking systems; 14 have the power to close, reconstitute, or take over low-performing schools; 16 have the authority to replace teachers or administrators; and 11 have the authority to revoke accreditation. In addition, Amrein and Berliner found that not only teachers, administrators, and schools were the targets of rewards and punishments, but also students.

High-Stakes Testing and Teachers' Philosophies of Education

A philosophy of education involves one's ideas about the purposes of education. It develops from conversations with others, such as students, colleagues, and the public,

as well as from authors one reads. These conversations tend to cause educators to reflect on their own ideas and consider other perspectives. A philosophy of education is constantly changing and evolving; it is never complete. Each conversation has an impact on one's philosophy of education. These conversations are the foundation for public education in this country. Teachers, principals, parents, students, and policymakers all come to the table and share their philosophies of education and form school communities. Discourse about philosophies of education is fundamental to the survival of public education in this country. A clear philosophy of education must exist to answer the questions of why and how education takes place. The lack of answers to these questions leads to constant school reform efforts that are usually short-lived and ineffective (Gunzenhauser, 2003).

The rush to accountability has had numerous effects on public schools, one of which involves the purpose of education. The impact is substantial because more emphasis is being placed on the scores achieved on high-stakes tests rather than on the achievement the scores represent. This has led to situations where accountability drives the curriculum and keeps educators from establishing their own visions and priorities for public education. Because of this, teachers may find themselves doing things that fall short of their vision of themselves as educators. These include spending extra time drilling students on practice tests and de-emphasizing or eliminating untested subject matter (Gunzenhauser, 2003).

Teachers' Response to High-Stakes Testing

According to Clarke and Gregory (2003), the jury is still out in regard to the impact of high-stakes tests on students, teachers, principals, schools, and school districts, and to whether or not these tests or such tests are appropriate for achieving excellence in American public schools. As a result, it becomes very important that policymakers, researchers, and others involve teachers and principals in this conversation; educators are the ones who must implement accountability policy.

One recent study (Pedulla et al., 2003) has provided strong empirical support for one side in this debate. The National Board on Educational Testing and Public Policy conducted a nationwide 2-year study that sought to ascertain teachers' attitudes and opinions about the impact of state-mandated testing programs on teaching and student learning (Pedulla et al.). The survey sample was designed to reflect the views of teachers in states with low, moderate, or high stakes attached to test results. Approximately 12,000 teachers were surveyed; however, only 4,200 teachers responded to the survey, for a response rate of 35%. Major areas surveyed included: (a) school climate, (b) pressure on teachers, (c) perceived value of the state test, (d) alignment of classroom practices with the state test, (e) impact on the content and mode of instruction, (f) test preparation and administration, (g) perceived unintended consequences, and (h) accountability and use of test results.

The majority of teachers among the survey respondents at each grade level, but particularly elementary teachers, indicated that state-mandated testing programs have led

them to teach in ways that contradict best practices. Regardless of the rewards and/or sanctions associated with the test, teachers felt as though state-mandated testing programs negatively impacted the quality of instruction that students received.

Across all states and grade levels, about 4 in 10 respondents indicated that teachers in their school could raise test scores without really improving learning. Three quarters of all teachers, regardless of state or grade level, found that the benefits of the testing programs were not worth the time and money involved. Teachers frequently reported that pressure to raise test scores encouraged them to emphasize instructional and assessment strategies that mirrored the content and format of state tests and to devote large amounts of classroom time to test preparation activities.

A majority of all teachers, however, were positive in their opinions of their state's curriculum standards, and the vast majority indicated that their district's curriculum was aligned with the state test. The majority of teachers across states and grade levels disagreed that the test was causing many students to drop out of high school or to be retained.

The findings from this study (Pedulla et al., 2003) are supported by other research that examined teachers' perceptions of the effects of high-stakes tests on teaching and on student learning. The majority of these studies indicated that teachers felt that high-stakes tests had a negative effect on curriculum, instruction, student learning, school climate, and teacher and student motivation. Most of the studies concluded that high-stakes tests had led to narrowing of curriculum, teaching to the tests, and spending less time on

content not covered by the tests (Adams & Karabenick, 2000; Barksdale-Ladd & Thomas, 2000; Brown, 1993; Cimbricz, 2002; Gordon & Reese, 1997; Grant, 2000; Hoffman, Assaf, & Paris, 2001; Jones & Egley, 2004; Kelley, Conley, & Kimball, 2000; Kubow & Debard, 2000; Mabry, Poole, Redmond, & Schultz, 2003; Parker, 1994; Rapp, 2002; Snow-Renner, 2001).

Although the majority of studies related to the impact of high-stakes tests on teaching and learning have concluded that high-stakes tests have a negative effect on teaching and student learning, studies by Schleisman (1999) and Snow-Renner (2001) have suggested that high-stakes tests do have some positive effects on teaching and student learning. Both of these studies explored teacher and principal perspectives on high-stakes tests. These studies found that high-stakes tests highlighted the needs of some students who otherwise might have slipped through the cracks, forced schools to address remediation needs of students, brought greater curricular coherence, generated information for school improvement, and helped teachers to guide teaching.

Wilson and Corbett (1991) conducted a study that investigated the effects of assessment on curriculum and instruction. Their study considered local educators' reactions to statewide minimum competency testing and instructional effects of implementing these tests in two states. The states examined were (a) Pennsylvania, a low-stakes situation with minor consequences for students; and (b) Maryland, a high-stakes situation where high school graduation depended on passing the test. Pennsylvania students were tested in Grades 3, 5, and 8. Maryland students were tested at the beginning

of Grade 9. Fieldwork was conducted at six sites in each state, and over 250 local educators were interviewed. A survey completed by 277 of 501 Pennsylvania districts and 23 of 24 Maryland districts provided additional information. Researchers found that the impact of the testing program was far greater in Maryland, the high-stakes situation, with respect to narrowing of the curriculum. There were marked differences in the responses between educators in the two states. In Pennsylvania, a low-stakes situation, approximately two thirds indicated that there was no change with respect to curriculum narrowing. In Maryland, one of seven respondents indicated no change, and two thirds of the teachers reported a moderate total change.

Another study (Kubow & DeBard, 2000) examined teachers' perceptions of proficiency testing in one Ohio suburban school district and found that the majority of the teachers felt that proficiency testing resulted in the school curriculum being aligned to fit the proficiencies and that math and science were given greater emphasis. Results from this study assisted educators in their efforts to better identify curricular weaknesses for each of the subject areas.

Principals' Response to High-Stakes Testing

There is extensive literature related to the effects of high-stakes tests on instruction. The proportion of research related to principals' perceptions of high-stakes tests is very limited, however, when compared to the research related to teachers' perceptions of high-stakes tests.

Two studies (Mitchell, 1996; Zellner & Jinkins, 2001) examined principals' perceptions of high-stakes tests on teaching and student learning. The primary purpose of both of these studies was to provide more information about the unintended consequences of high-stakes tests on teaching practices and student achievement. Descriptive quantitative and qualitative methods were used in the research designs of both of these studies. The researchers found that the use of high-stakes tests resulted in (a) a negative perception of school by teachers and students, (b) fewer opportunities for teachers to adjust instruction and curriculum to address student needs, (c) overemphasis by teachers on skills being assessed rather than integration of skills, and (d) overemphasis on basic skills.

Graham et al. (2002) conducted a survey of elementary school principals in Virginia. Surveys were sent to 1,167 principals; 547 surveys were returned. This study investigated the effects of high-stakes testing on elementary school art, music, and physical education. In the survey, principals were asked to report the amount of time that was allocated each week for specialists to teach art, music, and physical education during the academic year at each grade level in their schools. Over 83% of the principals reported no change from the previous school year in the amount of time allocated to art and physical education. More than 88% of the principals planned no changes for the next school year in the time allocated to art, music, or physical education as a means to allow more time on subjects being tested on state tests.

Reed, McDonough, Ross, and Robichaux (2001) conducted interviews with 26 principals in selected schools in south Florida. They addressed three questions: (a) To what degree does a school's standardized test grade influence a principal's sense of empowerment? (b) To what degree and in what ways is morale affected by high-stakes testing? and (c) What lessons do empowered principals have to share with others about the impact that empowerment has on the quality of teaching and learning in schools? Results showed that testing had a negative effect on principals' sense of empowerment and appeared to generate a pervasive fear of failure in lower performing schools. High performing schools appeared to use high-stakes testing in positive ways such as incentives to create student enrichment programs. Lower performing schools seemed to focus more on meeting students' basic needs. The study findings suggested that standardized test scores do have an impact on the types of teaching practices and opportunities for enrichment that are available for students. Study results also revealed that a great deal of attention is placed on test scores.

Acker-Hocevar and Touchton (2001) examined the perceptions of 10 principals from high poverty/minority schools regarding accountability measures placed upon their schools by the state. Data collection included field notes, observations, and interviews regarding principals' impressions of the state's accountability system. The study revealed that principals felt they faced continual pressure to improve student performance and meet state and district mandates while supporting teachers. The respondents felt

threatened by mandates from the top and reported losing teachers to other professions and schools due to poor teaching conditions.

Consequences of High-Stakes Testing

Over the past 2 decades, efforts to improve public education have increasingly focused on the use of state-mandated tests to ensure educational equity for all students. These tests are being used to measure student achievement, evaluate teacher competence, guide educators in placing students, and determine school quality (Darling-Hammond, 1991). These state-mandated tests generally focus on reading, writing, and math. States have implemented positive rewards for high levels of performance and negative consequences for low levels of performance. The negative consequences include retention in grade level, denial of diplomas, and publication of ratings for schools and districts.

Numerous research studies have investigated the consequence of high-stakes testing programs on schools, teachers, and students (Adams & Karabenick, 2000; Barksdale-Ladd & Thomas, 2000; Brown, 1993; Cimbricz, 2002; Gordon & Reese, 1997; Grant, 2000; Hoffman et al., 2001; Jones & Egley, 2004; Kelley et al., 2000; Kubow & Debard, 2000; Mabry et al., 2003; Parker, 1994; Rapp, 2002; Snow-Renner, 2001). Most of these studies have gathered information from teachers, administrators, students, and parents by using interviews, classroom observations, and surveys. These studies have yielded both positive and negative results; however, the preponderance of findings has

not been favorable toward the use of high-stakes tests as instruments to improve public education. Most of these studies have focused on classroom practices, teachers, and students (Abrams et al., 2003).

Positive Effects

On the positive side, there is evidence that high-stakes tests have led to school districts revisiting the district curriculum and forcing teachers to focus the curriculum to meet the needs of all students. For example, Schleisman (1999) conducted a study to provide an in-depth look at one school district's response to an externally mandated, high-stakes testing program in Minnesota. This study focused on middle and high schools only. In general, the educators in this district told about one of the most positive aspects of the adoption and implementation of Minnesota's basic standard to highlight the needs of some students who may have slipped through the cracks. Respondents also felt that the testing policy resulted in more attention being paid to helping ESL students increase their reading and writing proficiency. Another finding was that it forced the educators to focus more on reading in general for all students. Additional positive effects were changes in teachers' instructional practices. Vogler (2002) surveyed 257 teachers to determine if public release of student results on state-mandated high-stakes tests influenced teachers' instructional practices. The data showed a notable increase in the use of open-ended response questions, critical-thinking questions, problem-solving activities, writing assignments, and inquiry/investigation. Survey results also indicated that there was a decrease in the use of multiple-choice and true-false questions, textbook-based

assignments, and lectures.

Neutral and Negative Effects

Despite these positive findings, a large amount of the existing research on the impact of high-stakes tests on teaching and student learning describes neutral or negative effects. Many of these effects appear to diminish students' exposure to curriculum, which undermines the purpose of testing (Stecher, 2002).

Orlich (2003) analyzed a 4-year data set on student achievement from the required high-stakes test in Washington, the Washington Assessment of Student Learning (WASL). The study found no positive or negative effect on yearly student achievement as a consequence of longitudinal administration of the WASL. The findings of this study supported the findings of Amrein and Berliner (2002), who examined 18 states with severe consequences attached to their testing program to see if high-stakes testing programs were affecting student learning. These researchers found that in all but one analysis, student learning remained at the same level it was before the high-stakes testing policy was implemented. In some cases, student learning actually declined.

The majority of research studies related to state-mandated testing programs have addressed the effects on what was taught (Abrams et al., 2003). According to most of the findings, state-mandated high-stakes tests resulted in unintended negative effects such as narrowing of the curriculum, teachers adapting their teaching styles to match state tests, inflated test scores, cheating, increasing dropout rates, and inequities for minority students.

Narrowing of curriculum. According to Shepard (1991), as the political pressure for students to perform better on high-stakes tests increases, there is a tendency to teach content that is tested, to the exclusion of content that is not tested. Teachers are beginning to spend more time on tested subjects such as math and reading and are neglecting science and social studies. Instructional time is being taken away from areas such as art, music, and physical education to allow more time for the subjects that are tested by the state. There is a great deal of empirical research to support this assertion. For example, the National Board of Educational Testing and Public Policy study of teacher attitudes (Pedulla et al., 2003) indicated that state tests have a differential impact on what content gets emphasized depending on the level of the stakes. In high-stakes states, 43% of the responding teachers indicated that the amount of time they spent on instruction in tested areas had increased, while only 17% of teachers in low-stakes states increased instruction time in tested areas. The results from this study have been supported by other research studies (Gordon & Reese, 1997; Hoffman et al., 2001; Zellner & Jinkins, 2001).

Graham et al. (2002) conducted interviews with 360 educators in three states. Findings revealed that respondents were engaged in varying degrees of removing, emphasizing, or adding curriculum content in order to prepare for the state tests.

As the pressure for youngsters to perform well on high-stakes tests increases, so does the threat to reduce or eliminate art, music, and physical education programs so that teachers have more time in the day to focus on teaching and learning.

(Graham et al., p. 51)

Maurice and Karr-Kidwell (2003) used data from a field research project to enable teachers and instructional leaders to understand and overcome the problems associated with high-stakes testing, especially as they relate to the narrowing of the curriculum. A survey was sent to 48 teachers at a middle school in Texas; however, only 27 responded to the survey. The majority of the respondents registered negative feelings and opinions about standardized testing. More than 80% strongly agreed that they felt pressured by standardized testing, 96.3% agreed that their students also felt pressured, and 70% agreed that they taught to the test.

Adapting teaching styles to test formats. Due to the pressure being exerted on teachers to improve students' performance on high-stakes tests, some are adapting their teaching styles to make the format of classroom presentations more like that of the tests. For example, Pedulla et al. (2003) found that 51% of the teachers in high-stakes states, as compared to 29% of the teachers in low-stakes states, reported that their classroom tests were in the same format as the state test. In the same study, 76% of the high-stakes teachers reported that their state-mandated testing program had led them to teach in ways that contradicted their ideas about sound educational practice. For example, teachers reported that they were spending more time on tested content. They also reported that state tests influenced the frequency and manner in which they assessed students.

Inflated scores. According to Shepard (1991), political pressure and media attention attached to high-stakes test scores can lead to inflated test scores. This results in a false impression of student achievement. Klein, Hamilton, McCaffrey, and Stecher

(2000) examined the results of the Texas Assessment of Academic Skills (TAAS), the highest profile state-testing program and one that has recorded extraordinary recent gains in math and reading scores. To investigate whether the dramatic math and reading gains on TAAS represented actual academic progress, the researchers compared TAAS gains to score changes in Texas on the NAEP. Texas students did improve significantly more on the fourth-grade NAEP test than their counterparts nationally; however, the size of this gain was smaller than their gains on TAAS and was not present on the eighth-grade math test. The researchers concluded that the gains on TAAS were due to (a) students being coached to develop skills unique to that test, (b) narrowing of the curriculum to improve test scores, and (c) increases in activities that reduced the validity of the scores (e.g., extended time for completion of test items and cheating).

Amrein and Berliner (2002) examined test results in 18 states with severe consequences tied to their testing programs to see if these programs were affecting student learning. Evidence from their study indicated that student learning remained at the same level it had been before the policy was implemented. The study was supported by a later study conducted by Orlich (2003).

Cheating. Cheating is another negative reaction to high-stakes testing. There is a paucity of data in regard to incidences of cheating related to high-stakes tests; however, use of high-stakes can be expected to increase cheating due to political pressure. Cheating can take many forms, including providing test items in advance, suggesting revisions, making changes to the answer sheets before they are sent to the district office or the state,

leaving relevant material posted during test administration, and providing hints during the test (Hamilton et al., 2002). For example, in Chicago, a test audit was conducted due to a concern regarding the validity of the scores from Chicago public elementary schools on the Iowa Tests of Basic Skills. There were 2 schools selected from each of 20 administrative districts. Each school tested one seventh- and one eighth-grade class. Other schools were retested to obtain a geographical balance. In total, there were 23 suspect and 17 comparison schools. Among the 80 classrooms, 19 had score declines that met or exceeded the required cutpoints; 17 of these classrooms were in suspect schools. Educators in these schools were believed to have exceeded the publisher's time limits, and in some instances researchers detected that answer sheets had been altered (Perlman, 1985).

Shepard and Cutts-Dougherty (1991) conducted a study on the effects of testing on instruction and student learning. A total of 360 teachers in Grades 3, 5, and 6 in approximately 100 schools in two districts answered a questionnaire on test preparation/coaching practices and the effects on instruction. Teachers reported that they felt pressure from the district administration and the media to improve their students' scores, and that this led to extensive time being given to test preparation. While blatant cheating was rare, practices that would boost test scores, such as rephrasing of questions, were considered to occur more frequently.

Increasing dropout rates. Determining the impact of high-stakes tests on dropout rates is a very difficult task. Several factors, such as the end of social promotion, an

increase in the number of immigrants, and changes in graduation requirements, make it difficult to identify a single influence as the cause for students leaving school without graduating (Horn, 2003).

There is a growing body of literature related to this issue. The majority of the research studies and literature available relate high-stakes tests to an increase in the number of students dropping out of school. For example, Clarke, Haney, and Madaus (2000) examined how high-stakes tests affect dropout rates and high school completion rates. Their study examined the effects of minimum competency testing (MCT). The researchers found there was no MCT in half of the 10 states with the lowest dropout rates. The states with the highest dropout rates had MCT programs with standards set at least by the state. The researchers' conclusion was that high-stakes testing programs are linked to decreased rates of high school completion.

Marion and Sheinker (1999) reviewed empirical findings related to state-level MCT programs. They found evidence of unintended negative effects of these programs, including lack of transfer of higher order skills, narrowing of the curriculum to test content, corruptibility of high-stakes tests, testing time as time taken away from teaching, and increased dropout rates (especially for minority and low-achieving students).

In one of the most recent large-scale studies, Haney (2000) examined the impact of the TAAS on high school completion rates. Findings of that study suggested that the exit exam was associated with an increase in high school dropout rates, especially among minorities such as African Americans and Hispanics.

Effects on equity. It is important to note that the primary rationale for test-based accountability programs such as NCLB and EIA is to improve educational equity for all students in America's public schools. It is not clear, however, that these programs lead to more equal educational opportunities for students of color and low socioeconomic level. It has been argued that the negative effects of high-stakes testing programs appear to be greater for minority and low-income students than for high-performing students (Shepard, 1991). This argument is supported by a number of studies. For example, Coleman et al. (1966) conducted the Equality of Educational Opportunity Survey, a longitudinal study known as the Coleman Report. This national survey covered approximately 4,000 elementary and secondary schools. It compared the distribution of resources and opportunities among students of different races: blacks, Puerto Ricans, American Indians, Oriental Americans, and whites. The Coleman Report was significant in that it was a shift in research focus from inputs to results. Before this study, education reform focused solely on resources, such as facilities, teachers' salaries, textbooks, and supplies, under the assumption that better provision of these resources would fix the problems of public schools. The Coleman Report found American education to be unequal in most regions of the country. Nearly all white students were taught by white teachers, and most black students were taught by black teachers. The researchers concluded that the level of skill of the teachers might have contributed to this inequality through the matching of teachers and students. The results of this study led reformers to

focus on performance rather than increases in resources (Coleman et al., 1966; Firestone, Schorr, & Monfils, 2004).

Research regarding the effects of high-stakes testing on equity has been very limited. For example, Denoyer and White (1992) conducted a study that investigated the relationship between socioeconomic characteristics of Ohio school districts and district test performance. The findings showed that district performance ranking was largely due to family wealth and did not reflect school effectiveness or instructional quality. These researchers recommended that school districts not be ranked according to test scores, which are highly influenced by wealth, but on the basis of access to knowledge, district expectations, and the teaching environment. Darling-Hammond (1997) reached this conclusion:

Although the United States came sooner to the task of educating a wide range of students in public schools, it has yet to meet the challenge of providing equal access to quality education. What students have the opportunity to learn is typically a function of where they live and what their parents earn, and the color of their skin. (p. 264)

Professional Development

Over the years, the nature of staff development has changed. The amount of time required for teacher participation in staff development activities has increased, but there have been very few studies to explore the connection between staff development and

student achievement on standardized tests. With the current emphasis on accountability, building-level and central office administration are being pressured to show that inservice training changes teacher behavior and enhances student achievement.

With the advent of NCLB, state accountability has gained momentum. According to Holloway (2003), the focus of professional development needs to change. Schools should use results of their student assessments to identify their needs for professional development and then use this information to plan appropriate activities. Additionally, schools need to help teachers learn how to use student performance data to modify and target instruction in their classrooms.

Singh and McMillan (2002) conducted a study to identify effective staff development practices in schools with high test scores on state-mandated tests over a 2-year period. The participants were Algebra I and English teachers in two elementary, two middle, and two high schools. The results indicated that some practices identified by the teachers and principals as having contributed to higher test scores were the same for all three school levels. Some of these practices were (a) decentralization of staff development; (b) a strong, clear, in-depth focus on specific areas; (c) analysis of test scores to determine staff development needs; and (d) teacher input.

A relevant study by Turchi, Johnson, Owens, and Montgomery (2002) examined the impact of high-stakes tests on teachers' professional development to meet the incentives presented by the state's accountability system. Many of the teachers who had implemented changes in focus and teaching strategies due to professional development

reported that their students' behavior, attitudes, and learning had improved. Teachers reported a positive impact on student learning as a result of new teaching strategies they were using that assisted them in preventing students from falling through the cracks.

Stecher and Barron (2003) conducted a 2-year study of the impact of standards-based assessment on classroom practices in Kentucky. The study involved a survey of Kentucky teachers' classroom practices during the 1996-1997 and the 1997-1998 school years. A total of 479 teachers completed the survey, for an overall response rate of 54%. The researchers found that Kentucky teachers participated in professional development activities that were consistent with their state-mandated tests. For example, the researchers found that elementary teachers in Grades 4 and 5 who taught all subjects received more professional development on the subjects that were assessed in their grade level than on subjects that were not assessed. Additionally, the researchers found that teachers focused more on relevant content areas being assessed as a part of the state's accountability system than on nontested content.

The Principal as Instructional Leader

The escalating calls by the public, local school boards, and state and national officials for increased accountability have drastically changed the role of the principal. Principals have seen their responsibilities become increasingly difficult, to the point where they have become almost overwhelming. Principals have found themselves in the middle of an accountability storm. Both American society—which is conditioned by

instant gratification—and change experts expect immediate results from the latest reform efforts (Cooley & Shen, 2003).

Policymakers currently mandate accountability through student performance on state-mandated tests, assessment standards, and dissemination and publication of test results in the media. The consequences for not meeting student performance targets affect student graduation, district funding, and retention of principals. These consequences have placed increased pressures on principals to collaborate with teachers to ensure that goals for student learning are linked to effective strategies. The public's perception that public schools are on the decline, combined with these additional demands on principals, has resulted in a cry for more effective principal leadership to address the accountability requirements of the public and the local, state, and federal policymakers. This presents a problem for secondary school principals, who must find ways to improve student performance levels. The emphasis on accountability has resulted in additional pressure on the principals in their role as instructional leaders (Cooley & Shen, 2003).

According to Wiggins (1994), the role of the principal as instructional leader has been evolving over the last quarter century, and there have been many definitions for the term *instructional leadership* over the years. This term was popularized during the effective schools movement; however, in the effective schools model, leadership focused on instructional practice and its impact on student achievement. Under this model, the role of instructional leader includes (a) determining appropriate modes of instruction, (b) selecting the most effective materials to provide that instruction, (c) providing a model

for application of those materials, and (d) assessing whether the model is being implemented appropriately.

The role of the principal as instructional leader has been expanded over the years to include more focus on teacher growth and development. This change can be seen in the proliferation of literature that emphasizes the principal's role in creating a vision and establishing cultural norms for the school.

Cooley and Shen (2003) conducted a national study to investigate secondary school principals' perceptions of the status of the accountability movement and their professional responsibilities in the context of accountability. The data collected from a nationally representative sample of secondary school principals provided a portrait of some of the challenges facing principals in this environment of school reform. Three trends emerged from the secondary school principals' responses. First, even though several measures are used to assess school performance, testing is still the most dominant method used by states and local boards of education. Second, principals are working in very political environments. Third, principals are being called on to engage in leadership initiatives, such as instructional leadership, to improve teaching and learning.

Bossert, Dwyer, Rowan, and Lee (1982) asserted that principals can have a positive impact on student learning. The authors observed that principals can accomplish this through their manipulation of instructional factors (e.g., the amount of time students spend on learning, class size, or curriculum organization) and through school climate

factors (e.g., expectations for students, parental involvement, and school order and safety).

Summary

Chapter Two focused on a review of literature related to the evolution of accountability and the history of high-stakes testing. Literature regarding teachers' philosophies of education, standardized tests, teachers' and principals' beliefs, professional development, and principal leadership was also examined.

In reviewing the research related to the effects of high-stakes testing on teaching and learning, it is very clear that there is a vast amount of literature related to the topic. Numerous research studies have been conducted to examine the impact of state-mandated testing programs—particularly those with high-stakes tests—on districts, schools, teachers, and students (Abrams et al., 2003; Adams & Karabenick, 2000; Amrein, 2002; Barksdale-Ladd & Thomas, 2000; Bersola, 2002; Brown, 1993; Cooley & Shen, 2003; Gordon & Reese, 1997; Kaplan & Owings, 2001; Mitchell, 1996; Vogler, 2002; Williams, 2002; Young, 1996). The majority of these studies have gathered information from teachers, however, and very few studies have focused on principals' perceptions of the impact of high-stakes tests on teaching and learning. For that reason, this study included principals as participants and examined their perceptions of the impact of state-mandated high-stakes tests on teaching and student learning.

CHAPTER THREE

Methodology

The purpose of this chapter is to provide a description and rationale for the research design selected for this study. Specifically, this chapter describes the purpose of the study and the research questions, research design, participant selection, instrumentation, data collection, and data analysis procedures.

Purpose and Research Questions

The purpose of this study was to explore secondary school principals' perceptions of the value and impact of state-mandated high-stakes tests on content and mode of instruction in Tennessee. This research inquiry is framed and guided by the following research questions:

1. What are secondary school principals' perceptions of the value of state-mandated high-stakes tests?
2. What are secondary school principals' perceptions of the impact of state-mandated high-stakes tests on the content and mode of instruction?
3. Do secondary school principals' perceptions of the value and impact of state-mandated high-stakes tests on content and mode of instruction differ when examined by school level (e.g., middle/junior high and high school) and category (e.g., urban, suburban, and rural)?

Research Design

A survey research design was selected for this study. Survey research designs are ideal when measuring people's attitudes and opinions and when there is a large number of respondents. According to Creswell (2003), "The purpose of survey research is to generalize from a sample to a population so that inferences can be made about some characteristic, attitude, or behavior of the population" (p. 154). Since the purpose of this study was to explore secondary school principals' perceptions of the value and impact of state-mandated high-stakes tests on content and mode of instruction in Tennessee, the use of a survey research method of inquiry seemed the most appropriate method for this study.

Participant Selection

The entire population of 541 middle and high school principals from public schools in the state of Tennessee was invited to participate in this study. This number included 272 middle school principals and 269 high school principals. The principals included in the study population were obtained from a list of all public school middle and high school administrators from the Tennessee Department of Education Web site (Tennessee Department of Education, 2005).

Instrumentation

Data were obtained from the population by means of a mailed survey, a copy of which is included in Appendix E. Participants used a Likert scale for their responses to

most of the survey items utilized to determine secondary school principals' opinion levels related to the value and impact of state-mandated high-stakes tests on content and mode of instruction. The principals were asked to indicate whether they "strongly agree" (4), "agree" (3), "disagree" (2), or "strongly disagree" (1) with statements. Additionally, they were asked if the amount of time spent on activities "decreased a great deal" (1), "moderately decreased" (2), "stayed about the same" (3), "moderately increased" (4), or "increased a great deal" (5). Other items asking for demographic information had common response options.

The survey instrument used in the Pedulla, Abrams, Madaus, Russell, Ramos, and Miao (2003) research, the *Teacher Survey on the Impact of State-Mandated Testing Programs*, served as a model for the current study, with permission from the National Board on Educational Testing and Public Policy. The original 80-item survey instrument was used to elicit teachers' attitudes toward state testing programs. Most of the items in the survey were geared toward understanding the beliefs of teachers about the influence of their state's testing program on classroom instruction and student learning. The survey consisted of items in the form of questions or statements related to standards-based education reform. For most of the items, a Likert scale was used to assess teachers' opinions. In addition to these closed-format items, the questionnaire contained open-ended questions that allowed the teachers to write in responses about the impact of state-mandated testing on their instructional practices.

The Pedulla et al. (2003) survey instrument addressed the following areas:

- Information about state and district testing programs
- Climate
- Relationship of the state-mandated test to the state curriculum frameworks and standards
- Beliefs about teaching, learning, and assessment
- Classroom activities relating to instructional and testing practices
- Test preparation and administration
- Use and reporting of test results
- Professional development related to state-mandated tests
- Perceived effects of the state-mandated test. (Pedulla et al., 2003, p. 16)

For the current study, the 11 scales used in the original instrument were reduced to 4 scales: perceived value (see Figure 1), tested areas (see Table 1), noncore content (see Table 1), and classroom activities scales (see Table 1). The survey used in this study has 40 items. For the purposes of the current research, demographic information from the original survey was redesigned to incorporate information about secondary school principals rather than teachers.

According to Creswell (2003), “[Field testing] is important to establish the content validity of an instrument to improve the questions, format, and the scales” (p. 158). For the current study, the revised instrument was field tested by administering the survey to a group of 25 high school and middle school principals in February 2006. A

6. Overall, the benefits of the state-mandated testing program are worth the investment of time and money.
7. Media coverage of state-mandated test results accurately reflects the quality of education in my state.
8. Scores on the state-mandated test results accurately reflect the quality of education students have received.
9. The state-mandated test has brought much needed attention to education in my district.
10. The state-mandated test is an accurate measure of student achievement as a teacher's judgment.
11. The state-mandated test motivates previously unmotivated students to learn.
12. The state-mandated test measures high standards of achievement.
13. The state-mandated testing program is just another fad.
14. Media coverage of state-mandated testing issues has been unfair to teachers.
15. Media coverage of state-mandated testing issues adequately reflects the complexity of teaching.
16. Teachers in my school have found ways to raise test scores without really improving learning.
17. The state-mandated test is not an accurate measure of what students who are acquiring English as a second language know and can do.
18. Score differences from year to year on the state-mandated test reflect changes in the characteristics of students rather than changes in school effectiveness.

NOTE. From "Perceived Effects of State-Mandated Testing Programs on Teaching and Learning: Findings from a National Survey of Teachers," by J. Pedulla, L. Abrams, G. Madaus, M. Russell, M. Ramos, and J. Miao, 2003, p. 142. Chestnut Hill, MA: Boston College. Copyright by the Center for the Study of Testing, Evaluation, and Educational Policy. Adapted with permission of the authors.

Figure 1. Perceived Value Scale

Table 1. Tested Areas, Noncore Content, Classroom Activities Scales

19. In what ways, if any, has the amount of time you spent on each of the following activities changed in your school in order to prepare students for the state-mandated testing program?	Scale
Instruction in tested areas	Tested areas
Instruction in tested areas with high stakes attached (e.g., promotion, graduation, teacher rewards)	Tested areas
Parental contact	Tested areas
Instruction in areas not covered by the state-mandated test	Tested areas
Instruction in physical education	Noncore areas
Instruction in foreign language	Noncore areas
Instruction in industrial/vocational education	Noncore areas
Instruction in the fine arts	Noncore areas
Enrichment school assemblies (e.g., professional choral group performances)	Activities
Class trips (e.g., circus, amusement park)	Activities
Field trips (e.g., museum tour, hospital tour)	Activities
Student choice time (e.g., games, computer work)	Activities
Organized play (e.g., games with other classes)	Activities
Classroom enrichment activities (e.g., guest speakers)	Activities
Student performances (e.g., class plays)	Activities
Administrative school assemblies (e.g., awards ceremonies)	Activities
Student free time (e.g., recess, lunch)	Activities

NOTE. From “Perceived Effects of State-Mandated Testing Programs on Teaching and Learning: Findings from a National Survey of Teachers,” by J. Pedulla, L. Abrams, G. Madaus, M. Russell, M. Ramos, and J. Miao, 2003, p. 144. Chestnut Hill, MA: Boston College. Copyright 2003 by the Center for the Study of Testing, Evaluation, and Educational Policy. Adapted with permission of the authors.

total of 15 of these principals responded to the survey, for an overall response rate of 60%.

The purpose of this field test was to allow these principals who are experts in the field of education to comment on the flow and clarity of the survey, in questions and layout, from their perspective. According to Nardi (2006),

The best way of assessing whether the questionnaire flows, the instructions are adequate, the wording of the items and format are clear, and the survey takes a reasonable time to complete is to pilot test it—first with yourself and then with others. (pp. 95-96)

A Survey Modification Form was created and attached with the survey during the field test (see Appendix A). This form was included to give the principals an opportunity to comment on the adequacy of the items in the survey and to provide comments about the layout of the survey and the clarity of its questions and instructions. One change was made as a result of two principals' comments on the modification form. These principals commented that questions # 6 and #14 were the same. This typographical error was corrected in the final survey. The other 13 principals indicated that no changes were needed.

One of the most important characteristics of a measuring instrument is validity. It involves the appropriateness of the interpretations that are made based on the test results. There are four types of test validity: content validity, criterion-related validity, construct validity, and consequential validity (Gay & Airasian, 2003).

For the purposes of this research, content validity was used to establish validity of the survey instrument utilized in the study. Content validity is determined by expert judgment; there is no formula by which it can be computed or expressed quantitatively. Experts with content knowledge in the field covered by the instrument are asked to assess its content validity. These experts examine the process used to develop the instrument as well as the instrument itself. They then make a decision about how well the items represent the intended content.

The secondary school principals who participated in the field test of the survey are considered as experts in the field of education. They were used to determine the degree of content validity of the survey instrument. All of the principals who returned their surveys commented that the survey items adequately represent the topic. According to the principals involved in the field test, the current survey has a high degree of content validity. The principals involved in the pilot study were not included in the actual study because they had already seen the survey. Having them take the survey for a second time would taint the results (Nardi, 2006).

Reliability refers to the extent to which results can be replicated if the study is replicated (Merriam, 1998). According to Gay and Airasian (2003), reliability is expressed numerically, usually as a reliability coefficient that is obtained by using a correlation. A high reliability coefficient indicates high reliability. Cronbach's alpha was used in the original study to indicate the reliability of each scale (see Table 2). The value of Cronbach's alpha ranges between 0 and 1, with larger values indicating higher

Table 2. Scale Reliability

Scale	Scale reliability (Cronbach's alpha)
Perceived value	.79
Tested areas	.57
Noncore content	.83
Classroom activities	.91

Note. .57 is a less than acceptable level of reliability.

reliability. The reliability coefficient would be 1.00 if a test were perfectly reliable; however, no instrument is perfectly reliable.

Data Collection

Data were collected from the population by means of a mailed survey (see Appendix E). According to Nardi (2006), self-administered surveys are best designed for (a) measuring variables with numerous values or response categories that are too much to read to respondents in an interview or on the telephone, and (b) investigating attitudes and opinions that are not usually observable.

In this study, data collection was designed as a five-phase process. The first phase was the mailing of the prenotification letter to all members of the population 5 days prior to the mailing of the actual survey (see Appendix B). This letter notified the participants that an important survey would arrive in a few days and that their response would be greatly appreciated. The second phase was the mailing of a detailed cover letter (see

Appendix C) and the actual survey. Each survey had an identification number. This allowed the researcher to determine which surveys have been returned; however, confidentiality was maintained throughout the process. The third mailing consisted of a thank-you postcard 5 days after the mailing of the survey (see Appendix G). The fourth phase was the mailing of a second copy of the survey with (see Appendix E) a different cover letter (see Appendix H). This mailing was sent to nonrespondents 3 weeks after the initial mailing. Since a response rate of 60% was not achieved, then a fifth phase was initiated. This involved a phone call to all nonrespondents (Creswell, 2003; Dillman, 2000; Gay & Airasian, 2003).

Data Analysis

Analysis of the data in this study was performed using the Statistical Package for Social Sciences (SPSS). This software package offers a variety of statistical procedures for analyzing data. These include descriptive statistics, such as frequencies, percentages, and means. This package also offers inferential statistics, such as F-tests, t-tests, and chi square, as some of its features.

A response to research question 1, regarding middle and high school principals' perceptions of the value of state-mandated high-stakes tests, was generated using descriptive statistics. A mean and standard deviation were computed for all middle school principals' and high school principals' responses for each survey item. Means above 2.5 indicated agreement with an item where as means below 2.5 indicated disagreement. One

sample t-tests were run to determine if the mean for each item was significantly different from the mean score of 2.5.

The response to research question 2, regarding middle and high school principals' perceptions of the impact of state-mandated high-stakes tests on the content and mode of instruction, was generated using descriptive statistics. A mean and standard deviation were computed for each survey item. A mean score above 3.00 indicated an increase in the amount of instructional time spent in the area. One sample t-tests were run to determine if the mean of each item was significantly different from the mean score of 3.00.

The response to research question 3, regarding variance of middle and high school principals' perceptions of the value and impact of state-mandated high-stakes tests on the content and mode of instruction, was generated using descriptive statistics. Means and standard deviations were generated for each survey item by school level (e.g., middle and high school) and category (e.g., urban, suburban, and rural). Additionally, a Multivariate Analysis of Variance (MANOVA) was performed to determine if there are significant differences between middle and high school principals' responses. If the MANOVA was significant, Individual Analysis of Variances (ANOVAs) were performed to determine which items differed by school level and category.

CHAPTER FOUR

Research Findings

Introduction

The purpose of this study was to explore secondary school principals' perceptions of the value and impact of state-mandated high-stakes tests on content and mode of instruction. This research inquiry was framed and guided by the following research questions:

1. What are secondary school principals' perceptions of the value of state-mandated high-stakes tests?
2. How do secondary school principals perceive the impact of state-mandated high-stakes tests on the content and mode of instruction?
3. Do secondary school principals' perceptions of the value and impact of state-mandated high-stakes tests on content and mode of instruction differ when examined by school level (e.g., middle/junior high and high school) and category (e.g., urban, suburban, and rural)?

The survey used in the Pedulla et al. (2003) research, the *Teacher Survey on The Impact of State-Mandated Testing Programs*, was used as a model for the survey, *Secondary School Principals' Perceptions of the Value and Impact of State-Mandated Testing on Content and Mode of Instruction*, with permission from the National Board on Educational Testing and Public Policy. Surveys were sent to 541 middle and high school

principals across the state of Tennessee. Of this total, 272 were sent to middle school principals and 269 were sent to high school principals. The recipients returned a total of 310 surveys, for an overall response rate of 58%. Individual responses by the participants were analyzed and reported in composite form to preserve the anonymity of the participants. This chapter presents the demographic information and research findings of the statistical analysis of the data as they relate to the three research questions.

Analysis of Survey Demographics

All participants were asked to provide information about their schools and their years of administrative experience, gender, age, and race. Summaries of the demographic information are reported in Appendix F.

Question 1 on the demographics section asked principals which category—urban, suburban, or rural—best described their school. Although 302 principals returned their surveys, 3 principals (1%) did not respond to this question. Among those who did respond to this question, 58 principals (19.2%) indicated that their schools were urban, 79 principals (26.2%) reported that their schools were suburban, and 162 principals (53.6%) reported that their schools were rural (see Table 3).

Question 2 on the demographics section of the survey asked principals to report how their schools' performance on state-mandated tests compared to that of other schools in the state. Although 302 principals returned their surveys, 7 principals (2.3%) did not respond to this question. Among those who did respond to this question, 121 principals

Table 3. Category

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Urban	58	19.2	19.4	19.4
	Suburban	79	26.2	26.4	45.8
	Rural	162	53.6	54.2	100.0
	Total	299	99.0	100.0	
Missing	System	3	1.0		
Total		302	100.0		

(40.1%) reported their schools as performing above average, 156 principals (51.7%) indicated that their schools were performing at an average level, and 18 principals (6%) reported their schools as performing below average (see Table 4).

Question 3 asked principals to report their school's level: middle/junior high or high school. Although 302 principals returned their surveys, 3 principals (1%) did not respond to this question. Of the 299 principals who did respond to this question, 138 (45.7%) reported their schools as being middle/junior high schools, and 161 (53.3%) reported their schools as being high schools (see Table 5).

Question 4 required the principals to report the location of their schools within the state (i.e., east, middle, or west). Although 302 principals returned their surveys, 3 principals (1%) did not respond to this question. Among the 299 principals who did respond to this question, 113 (37.4%) reported their schools as being located in the eastern part of the state, 116 (38.4%) indicated that their schools were located in the

Table 4. School Performance

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Above Average	121	40.1	41.0	41.0
	Average	156	51.7	52.9	93.9
	Below Average	18	6.0	6.1	100.0
	Total	295	97.7	100.0	
Missing	System	7	2.3		
Total		302	100.0		

Table 5. School Level

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Middle School/Junior High	138	45.7	46.2	46.2
	High School	161	53.3	53.8	100.0
	Total	299	99.0	100.0	
Missing	System	3	1.0		
Total		302	100.0		

middle of the state, and 70 (23.2%) reported the western part of the state as the location of their schools (see Table 6).

Question 5 of the demographics section of the survey asked principals to report the size of their school as one of the following: 499 or fewer students, 500-999, 1000-1499, 1500-1999, or 2000 or more students. Although 302 principals returned their surveys, 3 principals (1%) did not respond to this question. Of those who did respond to this question, 92 principals (30.5%) reported that their schools had 499 or fewer students, 134 principals (44.4%) indicated that their schools had 500-999 students, and 46 principals (15.2%) reported their schools as having 1000-1499 students. In addition, 19 principals (6.3%) indicated that their schools had 1500-1999 students, and 8 principals (2.6%) reported that their school had 2000 or more students (see Table 7).

Question 20 asked principals to report their years of administrative experience. Of the 302 principals who returned surveys, 4 principals (1.3%) did not respond to this

Table 6. School Location

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	East	113	37.4	37.8	37.8
	Middle	116	38.4	38.8	76.6
	West	70	23.2	23.4	100.0
	Total	299	99.0	100.0	
Missing	System	3	1.0		
Total		302	100.0		

Table 7. School Size

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1-499	92	30.5	30.8	30.8
	500-999	134	44.4	44.8	75.6
	1000-1499	46	15.2	15.4	91.0
	1500-1999	19	6.3	6.4	97.3
	Over 2000	8	2.6	2.7	100.0
	Total	299	99.0	100.0	
Missing	System	3	1.0		
Total		302	100.0		

question. Among those who did respond, 4 principals (1.3%) reported 1 year of administrative experience, 20 principals (6.6%) reported 2-3 years of administrative experience, and 93 principals (30.8%) reported 4-8 years of administrative experience. In addition, 55 principals (18.2%) reported 9-12 years of administrative experience, 66 principals (21.9%) reported 13-20 years of administrative experience, and 60 principals (19.9%) reported 20 or more years of administrative experience (see Table 8).

Question 21 asked principals to report their gender. Of the 302 principals who returned surveys, 4 principals (1.3%) did not respond to this question. Among those who did respond, 77 principals (25.5%) reported their gender as being female, and 221 principals (73.2%) reported their gender as being male (see Table 9).

Question 22 asked principals to report their age. Of the 302 principals who returned their surveys, 4 (1.3%) did not respond to this question. Among those

Table 8. Participants' Years of Administrative Experience

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	4	1.3	1.3	1.3
	2-3	20	6.6	6.7	8.1
	4-8	93	30.8	31.2	39.3
	9-12	55	18.2	18.5	57.7
	13-20	66	21.9	22.1	79.9
	Over 20	60	19.9	20.1	100.0
	Total		298	98.7	100.0
Missing	System	4	1.3		
Total		302	100.0		

Table 9. Participants' Gender

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Female	77	25.5	25.8	25.8
	Male	221	73.2	74.2	100.0
	Total	298	98.7	100.0	
Missing	System	4	1.3		
Total		302	100.0		

who did respond, 1 principal (.3%) reported his age as being 20-30 years, 43 principals (14.2%) indicated their age as being 31-40, and 70 principals (23.2%) reported their age as 41-50 years. In addition, 157 principals (52%) reported their age as 51-60 years, and 27 principals (8.9%) indicated their age as being 60 or more years (see Table 10).

Question 23 asked principals to mark all of the categories that best described their race. The categories included African American, American Indian/Alaskan Native, Asian/ Pacific Islander, White, Hispanic, and Other. The principals were allowed to mark more than one category. Among the respondents, 34 principals (11.4%) reported their race as being African American, and 1 principal (.3%) reported his race as being American Indian or Alaskan Native. In addition, 263 (88.6%) reported their race as being White, and 1 principal (.3%) reported his race as being Hispanic (see Table 11).

Table 10. Participants' Age

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	20-30	1	.3	.3	.3
	31-40	43	14.2	14.4	14.8
	41-50	70	23.2	23.5	38.3
	51-60	157	52.0	52.7	90.9
	60+	27	8.9	9.1	100.0
	Total	298	98.7	100.0	
Missing	System	4	1.3		
Total		302	100.0		

Table 11. Participants' Race

	No		Yes	
	Count	%	Count	%
African American	263	88.6%	34	11.4%
American Indian or Alaskan Native	296	99.7%	1	.3%
Asian	297	100.0%	0	.0%
White	34	11.4%	263	88.6%
Pacific Islander	297	100.0%	0	.0%
Hispanic	296	99.7%	1	.3%
Other	297	100.0%	0	.0%

Analysis of Results

Research Question 1

What are secondary school principals' perceptions of the value of state-mandated high-stakes tests?

A response to research question 1, regarding middle and high school principals' perceptions of the value of state-mandated high-stakes tests, was generated using descriptive statistics. An overall mean and standard deviation for all middle and high school principals' responses were computed for each survey item related to research question 1.

Data from principals' responses to research question 1 are summarized in Table 12. Means above 2.5 indicate agreement with an item whereas means below 2.5 indicate disagreement. One sample t-tests were run to determine if the mean for each item was significantly different from the mean score of 2.5. Principals' responses to survey items 9, 14, 17, and 18 indicated that secondary school principals were in agreement that (a) State-mandated high-stakes tests are not an accurate measure of what ESL students know and can do, (b) media coverage of the results of state-mandated tests is unfair to teachers, (c) state-mandated high stakes tests have brought attention to education issues, and (d) score differences from year to year reflect changes in characteristics of students and not school effectiveness.

Table 12. Perceived Value of State-Mandated Test Scale

Item #	Survey item	N	Mean	Std. Deviation	p
17	Not an accurate measure of what ESL students know and can do	287	3.15	.583	.000*
14	Media coverage has been unfair to teachers	298	2.89	.598	.000*
18	Score differences from year to year reflect changes in characteristics in students not school effectiveness	295	2.81	.630	.000*
9	Brought attention to education issues	298	2.69	.680	.000*
6	Benefits worth time and money	297	2.56	.656	.122
16	Teachers raise state-mandated test scores without improving student learning	296	2.41	.631	.013*
13	Is a fad	295	2.22	.625	.000*
12	Measures high standards of achievement	296	2.18	.721	.000*
10	Accurate measure of student's achievement as a teacher's judgment	299	2.09	.677	.000*
8	Scores accurately reflect quality of education	299	2.06	.612	.000*
7	Media coverage adequately reflects quality of education	299	1.92	.589	.000*
15	Media coverage adequately reflects complexity of teaching	299	1.85	.612	.000*
11	Motivates previously unmotivated students	299	1.83	.668	.000*
	Valid N (listwise)	274			

Note. * denotes significant difference at $p < .05$ by the one sample t-test.

Principals' responses to survey items 7, 8, 10, 11, 12, 13, 15 and 16 indicated that principals did not agree that (a) media coverage of the results of state-mandated high-stakes tests adequately reflects the quality of education, (b) scores accurately reflect quality of education, (c) state-mandated high-stakes tests are as accurate a measure of student achievement as a teacher's judgment, (d) high-stakes tests motivate unmotivated students, (e) state mandated tests measure high standards of achievement, (f) high-stakes tests are a fad, (g) media coverage adequately reflects the complexity of teaching, and (h) teachers raise state-mandated test scores without improving student learning. See Table 12 for the specific responses.

Research Question 2

How do secondary school principals perceive the impact of state-mandated high-stakes tests on the content and mode of instruction?

A response to research question 2, regarding middle and high school principals' perceptions of the impact of state-mandated high-stakes tests on the content and mode of instruction, was generated using descriptive statistics. An overall mean and standard deviation were computed for each survey item.

To address research question 2, the survey items asked principals whether content and mode of instructional practices were being influenced by state-mandated tests. The following discussion is based on the principals' responses to survey item 19. The items composing question 19 were combined to form three scales: (a) Impact on Tested Subject Areas, (b) Impact on Noncore Subject Areas, and (c) Impact on Student and Classroom

and Student Activities. Item 19 presented principals with various content areas and asked, “In what ways, if any, has the amount of time spent on each of the following changed in your school in an effort to prepare students for state-mandated tests?” Principals selected their responses on a range from “decreased a great deal” (1) to “increased a great deal” (5).

Data regarding secondary school principals’ perceptions of the impact of state-mandated high-stakes tests on tested subject areas are presented in Table 13. The mean scores ranged from 2.28 to 4.32. A mean score above 3.00 indicated an increase in the amount of instructional time spent in the area. One sample t-tests were run to determine if the mean of each item was significantly different from the mean score of 3.00. Principals’ responses indicated an increase in the amount of time spent on subject areas that were tested. In contrast to instructional time in tested areas with high stakes attached, principals reported that the amount of instruction in areas not covered by state-mandated tests decreased. Additionally, principals reported that parental contact had increased somewhat as a result of state-mandated tests.

Table 14 presents the mean scores for the four items that form the Noncore Subject Area Scale (i.e., foreign language, industrial/vocational education, physical education, and fine arts). The mean scores for these items ranged from 2.73 to 2.93. Higher mean scores, greater than 3.00, indicated an increase in the amount of time devoted to an activity. One sample t-tests were run to determine if the mean of each item was significantly different from the mean score of 3.00. Overall, principals reported that

Table 13. Tested Areas Scale

	N	Mean	Std. deviation	p
Tested areas	299	4.32	.735	.000*
Tested areas with high stakes attached	293	4.20	.735	.000*
Parental contact	299	3.68	.750	.000*
Tested areas without high stakes attached	294	2.86	.924	.008*
Areas not covered by state-mandated tests	299	2.28	.883	.000*
Valid N (listwise)	292			

Note. * denotes significant difference at $p < .05$ by the one sample t-test.

Table 14. Noncore Areas Scale

	N	Mean	Std. deviation	p
Foreign language	276	2.93	.690	.117
Industrial/vocational education	279	2.84	.771	.001*
Physical education	298	2.78	.648	.000*
Fine arts	296	2.73	.722	.000*
Valid N (listwise)	270			

Note. * denotes significant difference at $p < .05$ by the one sample t-test.

the amount of time spent on instruction in fine arts, physical education, and industrial/vocational education moderately declined.

Data regarding principals' perceptions of the impact of state-mandated high-stakes tests on items listed in the Classroom and Student Activities Scale (i.e., assemblies, class trips, and field trips) are presented in Table 15. The mean scores for these items ranged from 2.12 to 2.66. Higher mean scores, greater than 3.00, indicated an increase in the amount of time devoted to an activity. One sample t-tests were run to determine if the mean of each item was significantly different from the mean score of 3.00. Across all of the items in this scale, principals indicated a decrease in the amount of time spent on these activities. Principals reported that instructional activities related to classroom enrichment seemed to decrease the least ($M = 2.66$), while class trips decreased the most ($M = 2.12$).

Research Question 3

Do secondary school principals' perceptions of the value and impact of state-mandated high-stakes tests on content and mode of instruction differ when examined by school level (e.g., middle and high school) and category (e.g., urban, suburban, and rural)?

The response to research question 3, regarding variance of middle and high school principals' perceptions of the value and impact of state-mandated high-stakes tests on the content and mode of instruction, when examined by school level and category, was generated using descriptive statistics. Means were generated for each survey item by

Table 15. Activities Scale

	N	Mean	Std. deviation	p
Classroom enrichment activities	299	2.66	.785	.000*
Student performances	299	2.59	.743	.000*
Administrative school assemblies	299	2.56	.802	.000*
Student free time	295	2.43	.734	.000*
Field trips	299	2.34	.748	.000*
Enrichment school assemblies	299	2.30	.828	.000*
Student choice time	297	2.24	.743	.000*
Organized play	297	2.22	.797	.000*
Class trips	297	2.12	.827	.000*
Valid N (listwise)	292			

Note. * denotes significant difference at $p < .05$ by the one sample t-test.

school level and school category.

Inferential statistics were also used to address research question 3. Table 16 presents the results. A MANOVA was run to explore differences in secondary school principals' perceptions of the value of state-mandated high-stakes tests when examined by school level and category. Significant differences in levels were found ($p = .006$). This indicates that at least one of the items significantly differed by level. To determine which individual items differed by level, individual ANOVAs were run. Results from the individual ANOVAs are presented in Table 17. Means for each item are presented in Table 18 by school level. Differences were found in the way principals responded to

Table 16. MANOVA Results by Level and Category on the Perceived Value Scale

Effect	Wilks' lambda	F	Hypothesis df	Error df	Sig.
Intercept	.007	2926.970(a)	13.000	256.000	.000
Category	.904	1.021(a)	26.000	512.000	.436
Level	.895	2.320(a)	13.000	256.000	.006
Category & level	.929	.744(a)	26.000	512.000	.818

Table 17. ANOVA for Perceived Value of State-Mandated Test Scale Tests of Between-Subjects Effects

Item #	Dependent variable	Type III sum of squares	df	Mean square	F	Sig.
6	Benefits worth time and money	.018	1	.018	.043	.836
7	Media coverage adequately reflects quality of education	.240	1	.240	.674	.412
8	Scores accurately reflect quality of education	.136	1	.136	.358	.550
9	Brought attention to education issues	.193	1	.193	.423	.516
10	Accurate measure of student achievement as a teacher's judgment	.460	1	.460	.998	.319
11	Motivates previously unmotivated students	4.198	1	4.198	10.150	.002*
12	Measures high standards of achievement	.457	1	.457	.880	.349
13	Is a fad	2.815	1	2.815	7.810	.006*
14	Media coverage has been unfair to teachers	.284	1	.284	.789	.375
15	Media coverage adequately reflects complexity of teaching	.943	1	.943	2.645	.105
16	Teachers raise state-mandated test scores without improving student learning	.059	1	.059	.144	.704
17	Not an accurate measure of what ESL students know and can do	2.186	1	2.186	6.532	.011*
18	Score differences from year to year reflect changes in characteristics in students not school effectiveness	.101	1	.101	.243	.622

Note. * denotes significant difference at $p < .05$.

Table 18. Perceived Value Scale Means by Level

Item #		Level			
		Middle School/Junior High		High School	
		Mean	Std. Error	Mean	Std. Error
6	Benefits worth time and money	2.427	.061	2.445	.062
7	Media coverage adequately reflects quality of education	3.123	.056	3.057	.057
8	Scores accurately reflect quality of education	2.971	.058	2.922	.059
9	Brought attention to education issues	2.303	.064	2.244	.064
10	Accurate measure of student's achievement as a teacher's judgment	2.964	.064	2.874	.065
11*	Motivates previously unmotivated students	3.353	.061	3.079	.061
12	Measures high standards of achievement	2.777	.068	2.867	.069
13*	Is a fad	2.916	.057	2.691	.057
14	Media coverage has been unfair to teachers	2.050	.056	2.122	.057
15	Media coverage adequately reflects complexity of teaching	3.215	.056	3.085	.057
16	Teachers raise state-mandated test scores without improving student learning	2.605	.060	2.573	.061
17*	Not an accurate measure of what ESL students know and can do	1.723	.054	1.922	.055
18	Score differences from year to year reflect changes in characteristics in students not school effectiveness	2.183	.061	2.226	.061

Note. * denotes items that differ significantly at $p < .05$.

three items: (a) motivates previously unmotivated students ($p = .002$), (b) is a fad ($p = .006$), and (c) not an accurate measure of what ESL students know and can do ($p = .011$).

To explore how these responses differed, the means were examined (see Table 19). The lower the mean score, the more participants disagreed with the item. Middle school principals indicated greater disagreement than high school principals did with this statement: “State-mandated high-stakes tests motivate previously unmotivated students.” High school principals’ responses to all of the items were higher, indicating more agreement, except for this statement: “State-mandated high-stakes tests are not an accurate measure of what ESL students know and can do.”

A MANOVA was run to explore the differences in secondary school principals’ perceptions of the impact of state-mandated tests on content and mode of instruction when examined by school level (e.g., middle and high school) and category (e.g., urban, suburban, and rural) (see Table 20). Significant differences were found by level ($p < .001$) in the way principals responded to five items on the Tested Areas Scale.

Individual ANOVAs were run to examine how individual items differed by level in the way principals responded to the items on the Tested Areas Scale. Results from the ANOVAs are presented in Table 21. Table 22 presents the differences that were found in the way principals responded to (a) areas not covered by state-mandated tests ($p < .001$) and (b) tested areas without high stakes attached ($p = .028$). The mean scores ranged from $M = 1.930$ to $M = 2.940$. Mean scores above $M = 3.00$ indicate an increase in

Table 19. Means for Differences in Perceived Value Scale

Dependent variable	Level	Mean	Std. Error
Motivates previously unmotivated students	Middle school/junior high	1.647	.061
	High school	1.921	.061
Is a fad	Middle school//junior high	2.084	.057
	High school	2.309	.057
Not an accurate measure of what ESL students know and can do	Middle school/junior high	3.277	.054
	High school	3.078	.055

Table 20. MANOVA Results by Level and Category on Tested Areas Scale

Wilks' lambda

Effect	Value	F	Hypothesis df	Error df	Sig.
Intercept	.013	4196.096(a)	5.000	282.000	.000
Category	.977	.648(a)	10.000	564.000	.773
Level	.859	9.278(a)	5.000	282.000	.000*
Category & level	.957	1.244(a)	10.000	564.000	.260

Note. * denotes significant difference at $p < .05$.

Table 21. ANOVA for Tested Areas Scale by Level

Tests of Between-Subjects Effects

Dependent variable	Type III sum of squares	df	Mean square	F	Sig.
Tested areas	.466	1	.466	.846	.359
Areas not covered by state-mandated tests	24.939	1	24.939	36.543	.000*
Tested areas with high stakes attached	1.643	1	1.643	3.038	.082
Tested areas without high stakes attached	4.116	1	4.116	4.881	.028*
Parental contact	.509	1	.509	.954	.330

Note. * denotes significant difference at $p < .05$.

Table 22. Means for Differences in Tested Areas Scale

Dependent variable	Level	Mean	Std. error
Areas not covered by state-mandated tests	Middle school/junior high	1.930	.076
	High school	2.575	.075
Tested areas without high stakes attached	Middle school/junior high	2.678	.084
	High school	2.940	.084

instructional time, and lower ones indicate a decrease. High school principals' mean scores were higher, indicating that they spent more time on areas not covered by state-mandated tests and tested areas without high stakes. Both middle and high school principals' mean scores were less than $M = 3.00$, however, indicating a decrease in the amount of time spent in these areas.

A MANOVA was run to explore the differences in secondary school principals' perceptions of the impact of state-mandated tests on noncore areas, when examined by school level (e.g., middle and high school) and category (e.g., urban, suburban, and rural) (see Table 23). Significant differences were found by level ($p < .001$).

Individual ANOVAs found level differences in the way principals responded to four items on the Noncore Areas Scale. The differences were found in the way principals responded to (a) fine arts ($p < .001$), (b) physical education ($p = .004$), (c) foreign language ($p < .001$), and (d) industrial/vocational education ($p < .001$). Results from the

Table 23. MANOVA Results by Level and Category on Noncore Scale

Effect		Value	F	Hypothesis df	Error df	Sig.
Intercept	Wilks' lambda	.044	1424.954(a)	4.000	261.000	.000
Category	Wilks' lambda	.942	1.994(a)	8.000	522.000	.045
Level	Wilks' lambda	.891	7.946(a)	4.000	261.000	.000*
Category & level	Wilks' lambda	.931	2.380(a)	8.000	522.000	.016

Note. * denotes significant difference at $p < .05$.

ANOVAs are presented in Table 24. To explore how those differed, the means were examined (Table 25). A high mean score ($M > 3.00$) indicated that more instructional time was devoted to this area. Overall, the means for middle and high school principals' responses indicated that there had been a slight decrease in the amount of instructional time related to these areas. The mean scores for high school principals' responses, however, were slightly higher than the mean scores for middle school principals' responses in all areas that differed significantly.

Table 24. ANOVA for Noncore Areas Scale Tests of Between-Subjects Effects

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.
Category	Fine arts	2.108	2	1.054	2.158	.118
	Physical education	.610	2	.305	.737	.480
	Foreign language	.987	2	.494	1.111	.331
	Industrial/vocational education	1.553	2	.776	1.407	.247
Level	Fine arts	10.672	1	10.672	21.851	.000*
	Physical education	3.464	1	3.464	8.367	.004*
	Foreign language	5.621	1	5.621	12.648	.000*
	Industrial/vocational education	13.160	1	13.160	23.853	.000*
Category & level	Fine arts	.896	2	.448	.917	.401
	Physical education	.906	2	.453	1.095	.336
	Foreign language	2.793	2	1.397	3.142	.045*
	Industrial/vocational education	.641	2	.321	.581	.560

Note. * denotes significant difference at $p < .05$.

Table 25. Means for Differences in Noncore Areas Scale

Dependent variable	Level	Mean	Std. error
Fine arts	Middle school/junior high	2.467	.068
	High school	2.903	.063
Physical education	Middle school/junior high	2.611	.063
	High school	2.859	.058
Foreign language	Middle school/junior high	2.743	.065
	High school	3.060	.060
Industrial/vocational education	Middle school/junior high	2.532	.073
	High school	3.016	.067

A significant interaction between category and level was found ($p = .016$). However, when examined by category and level (see Table 26), only the principals' responses to the amount of time spent on foreign language were slightly significant. Overall, the responses indicated that there had been a slight decrease in the amount of time spent in foreign language courses. Further examination revealed that in the urban, suburban, and rural categories, high school principals' responses were higher than middle school principals' responses (see Figure 2). This indicated that in a high-stakes testing environment, high schools were spending more time than middle schools were on foreign language courses. Additionally, rural high school principals reported that they spent more time on foreign language courses than urban or suburban principals reported. In contrast, middle school principals reported that they spent less time on foreign language courses, as a result of state-mandated tests, than high school principals reported. The mean score

Table 26. Means for Differences in Foreign Language by Category and Level

Dependent variable	Category	Level	Mean	Std. error
Foreign language	Urban	Middle school/junior high	2.852	.128
		High school	3.037	.128
	Suburban	Middle school/junior high	2.857	.113
		High school	3.026	.108
	Rural	Middle school/junior high	2.521	.096
		High school	3.116	.068

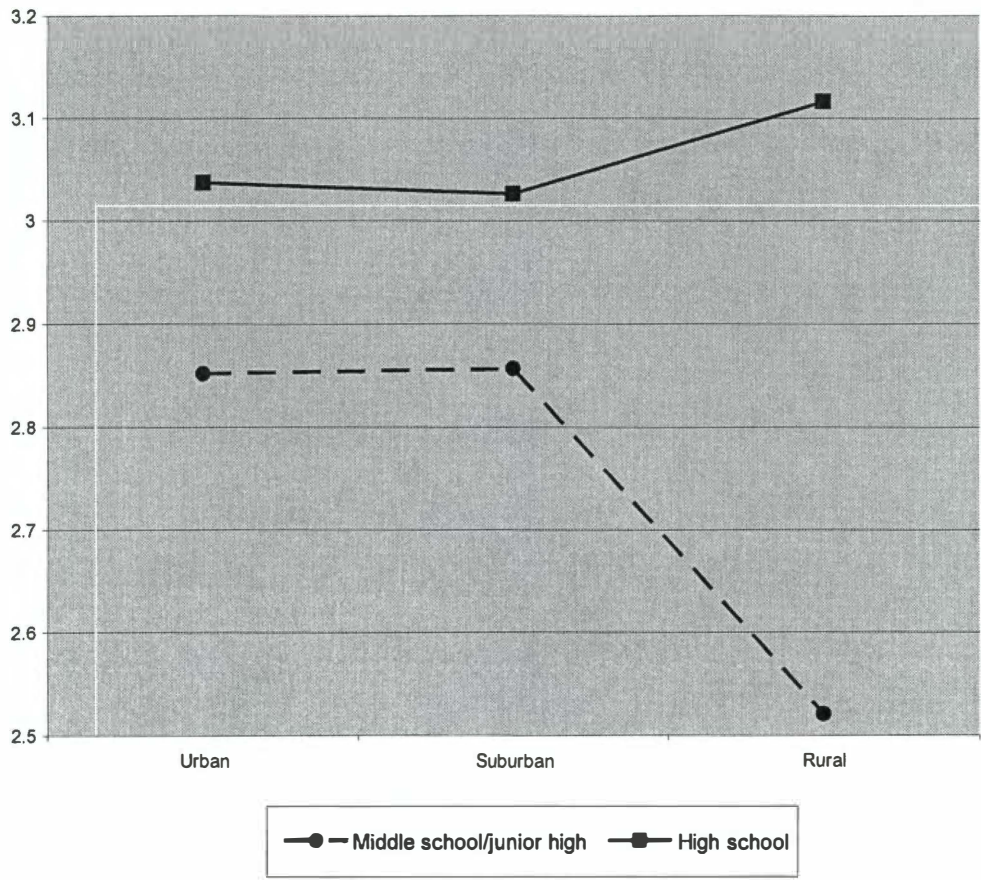


Figure 2. Foreign Language

for middle school principals in rural schools was lower than the mean scores for urban or suburban middle school principals.

A MANOVA was run to explore the differences in secondary school principals' perceptions of the impact of state-mandated tests on classroom and student activities, when examined by school level (e.g., middle and high school) and category (e.g., urban, suburban, and rural) (see Table 27). Significant differences were found by level ($p < .024$).

Individual ANOVAs found level differences in the way principals responded to four items on the Activities Scale. The differences were found in the way principals responded to (a) student free time ($p = .010$), (b) student choice time ($p = .008$), (c) organized play ($p = .007$), and (d) student performance ($p = .003$). Results from the ANOVAs are presented in Table 28. To explore how those differed, the means were examined (Table 29). A high mean score ($M > 3.00$) indicated that more instructional time was devoted to this area. Overall, the means for middle and high school principals'

Table 27. MANOVA Results by Level and Category on Activities Scale

Effect	Wilks' Lambda	F	Hypothesis df	Error df	Sig.
Intercept	.053	555.989	9.000	278.000	.000
Category	.963	.592	18.000	556.000	.906
Level	.934	2.174	9.000	278.000	.024*
Category & level	.933	1.087	18.000	556.000	.361

Note. * denotes significant difference at $p < .05$.

Table 28. ANOVA for Activities Scale by Level Tests of Between-Subjects Effects

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.
Level	Student free time	3.607	1	3.607	6.803	.010*
	Field trips	.225	1	.225	.395	.530
	Class trips	.308	1	.308	.446	.505
	Student choice time	3.884	1	3.884	7.081	.008*
	Organized play	4.622	1	4.622	7.340	.007*
	Enrichment school assemblies	.734	1	.734	1.064	.303
	Administrative school assemblies	.005	1	.005	.008	.927
	Classroom enrichment activities	.611	1	.611	1.003	.317
	Student performance	4.803	1	4.803	8.866	.003*

Note. * denotes significant difference at $p < .05$.

Table 29. Activities Scale Means by Level

	Level			
	Middle School/Junior High		High School	
	Mean	Std. Error	Mean	Std. Error
Student free time	2.291	.066	2.535	.066
Student choice time	2.091	.067	2.344	.067
Organized play	2.047	.072	2.323	.072
Student performance	2.411	.067	2.692	.067

responses indicated that there had been a decrease in the amount of instructional time related to classroom and student activities. The mean scores for high school principals' responses, however, were slightly higher than the mean scores for middle school principals' responses in all areas.

Summary

Chapter Four has presented the findings of this study, including an analysis of participant demographics and responses to the survey items related to the three research questions. Participants responded to the five demographic questions in common: category of school, school performance, school level, location, and size of their school. Additional demographic information was also discussed.

Descriptive statistics were used to describe secondary school principals' responses to research question 1: What are secondary school principals' perceptions of the value of state-mandated high-stakes tests? Means and standard deviations were computed for principals' responses to each survey item. The findings revealed that, overall, secondary school principals were in agreement that (a) State-mandated high-stakes tests are not an accurate measure of what ESL students know and can do, (b) media coverage of the results of state-mandated tests is unfair to teachers, (c) state-mandated high stakes tests have brought attention to education issues, and (d) score differences from year to year reflect changes in characteristics of students and not school effectiveness. In contrast to this, principals indicated that they disagreed with the

statements that (a) media coverage of the results of state-mandated high-stakes tests adequately reflects the quality of education, (b) scores accurately reflect quality of education, (c) state-mandated high-stakes tests are as accurate a measure of student achievement as a teacher's judgment, (d) high-stakes tests motivate unmotivated students, (e) state mandated tests measure high standards of achievement, (f) high-stakes tests are a fad, (g) media coverage adequately reflects the complexity of teaching, and (h) teachers raise state-mandated test scores without improving student learning.

Descriptive statistics were computed to describe principals' responses to research question 2: How do secondary school principals perceive the impact of state-mandated high-stakes tests on the content and mode of instruction? Means and standard deviations were computed for principals' responses to each item in the tested areas scale, noncore areas scale, and the activities areas scale. Findings revealed that, overall, principals perceived that their schools had increased the amount of time spent in subject areas with high-stakes tests. Additionally, principals indicated that the amount of parental contact time had been increased due to high-stakes tests. In contrast to an increase in the amount of time spent on tested subjects and an increase in parental contact, principals reported a decrease in the amount of time spent in all noncore subject areas (i.e., foreign language, industrial/vocational education, physical education, and fine arts). Additionally, principals reported a reduction in the amount of time spent on classroom and student activities (i.e., class trips, organized play, etc.).

Finally, descriptive and inferential statistics were performed to address secondary school principals' responses to research question 3: Do secondary school principals' perceptions of the value and impact of state-mandated high-stakes tests on content and mode of instruction differ when examined by school level and category? A MANOVA was run to determine if there were any differences in secondary school principals' responses. Significant differences were found in principals' responses when examined by school level. ANOVAs were run to explore the specific differences in middle and high school principals' responses. Results indicated that high school principals agreed more than middle school principals did that (a) state-mandated tests motivate previously unmotivated students, (b) state-mandated high-stakes tests are a fad, and (c) state-mandated high-stakes tests are not an accurate measure of what ESL students know and can do. Additionally, high school principals across all categories (i.e., urban, suburban, and rural) indicated that their schools spent more time on areas not covered on the state-mandated tests, while middle school principals did not indicate this. Results also revealed that high schools spent more time on foreign language than middle schools did.

Principals' responses were also examined by category. No significant differences were found by category.

CHAPTER FIVE

Summary, Findings, Conclusions, Discussion, and Recommendations

Summary

Since the publication of *A Nation at Risk* in 1984, the debate over high-stakes testing programs has escalated in the education, political, and research communities. The emphasis on high-stakes testing has emerged during what might be called the accountability era in education. This era came about because policymakers at the federal level were beginning to view America's public schools as failing due to a decline in student performance on standardized tests, poor academic performance, and a loss of economic competitiveness by the nation. All of these factors were seen as possible threats to the nation's existence. The authors of *A Nation at Risk* wrote:

Our Nation is at risk. Our once unchallenged preeminence in commerce, industry, science, and technological innovation is being overtaken by competitors throughout the world. This report is concerned with only one of the many causes and dimensions of the problem, but it is the one that undergirds American prosperity, security, civility. We report to the American people that while we can take justifiable pride in what our schools and colleges have historically accomplished and contributed to the United States and the well-being of its people, the educational foundations of our society are presently being eroded by a

rising tide of mediocrity that threatens our very future as a Nation and a people.

(National Commission on Excellence in Education, 1983, p. 1)

Because public schools are seen by the public and policymakers at the national, state, and local level as being in a state of decline, a great deal of legislation has been enacted to improve the effectiveness of America's public schools and to improve student learning. At the heart of this legislation, which includes the No Child Left Behind Act of 2001 at the federal level and other legislation at the state and local levels, is high-stakes testing. According to Gulek (2003),

Until the enactment of the No Child Left Behind Act, states had some leeway in determining whether to attach high stakes to the test results. However, with the passage of this act, every state-mandated testing program has become high stakes for schools and districts. (p. 42)

Many people who are involved in public schools argue that high-stakes testing undermines the main purposes of education. More emphasis is being placed on scores achieved on high-stakes tests rather than on the achievement the scores represent. This has led to situations where high-stakes tests are driving the curriculum, forcing teachers to teach in ways that contradict best practice. According to Pedulla et al. (2003), 7 out of 10 teachers in high-stakes states reported that state-mandated high-stakes tests have led them to teach in ways that contradict sound educational practice. This includes aligning their curriculum and assessments to mirror the state-mandated tests and devoting a large amount of instruction time to test preparation.

Even though there is a great deal of debate regarding the use of high-stakes tests to improve student performance and teacher effectiveness, those who do the real work, the principals and teachers at the building level, have been given very little say in the matter. Therefore, secondary school principals were included as participants in this study.

Purpose and Research Questions

The purpose of this study was to explore secondary school principals' perceptions of the value and impact of state-mandated high-stakes tests on content and mode of instruction in Tennessee. This research inquiry was framed and guided by the following research questions:

1. What are secondary school principals' perceptions of the value of state-mandated high-stakes tests?
2. What are secondary school principals' perceptions of the impact of state-mandated high-stakes tests on the content and mode of instruction?
3. Do secondary school principals' perceptions of the value and impact of state-mandated high-stakes tests on content and mode of instruction differ when examined by school level (e.g., middle and high school) and category (e.g., urban, suburban, and rural)?

Since the study involved the measurement of 541 secondary school principals' perceptions of the impact of high-stakes tests on content and mode of instruction, a

survey was used to gather data. All of the participants in the study were principals of secondary public schools comprised of some combination of Grades 6 through 12.

Summary of Findings

The following summary is made in response to the three research questions in this study. The first question asked: What are secondary school principals' perceptions of the value of state-mandated high-stakes tests? From an analysis of the data, the following issues were revealed: (a) State-mandated high-stakes tests are not an accurate measure of what ESL students know and can do, (b) media coverage of the results of state-mandated tests is unfair to teachers, (c) state-mandated high stakes tests have brought attention to education issues, and (d) score differences from year to year reflect changes in characteristics of students and not school effectiveness. Secondary school principals reported agreement related to these issues. The principals involved in this study disagreed with the following statements: (a) media coverage of the results of state-mandated high-stakes tests adequately reflects the quality of education, (b) scores accurately reflect quality of education, (c) state-mandated high-stakes tests are as accurate a measure of student achievement as a teacher's judgment, (d) high-stakes tests motivate unmotivated students, (e) state mandated tests measure high standards of achievement, (f) high-stakes tests are a fad, (g) media coverage adequately reflects the complexity of teaching, and (h) teachers raise state-mandated test scores without improving student learning.

The second research question asked: What are secondary school principals' perceptions of the impact of state-mandated high-stakes tests on the content and mode of instruction? To address this question, principals were asked if the amount of time spent in the following activities changed in their schools in order to prepare students for state-mandated testing: (1) tested subject areas, (2) noncore subject areas, and (3) classroom and student activities. Principals were asked to respond on a Likert scale that ranged from "decreased a great deal" (1) to "increased a great deal" (5). Principals' responses indicated an increase in the amount of time spent on subject areas that were tested. In contrast to the increase in the amount of instructional time spent in tested areas, principals reported a decrease in the amount of instructional time in areas not covered by state-mandated tests.

Principals' responses to the amount of time spent in noncore areas indicated a decrease in the amount of time spent in these areas. Principals reported that instruction in fine arts, physical education, industrial/vocational education, and foreign language moderately declined. Moreover, principals reported foreign language as having the smallest decrease in time, while fine arts decreased the most.

Principals' responses to items on the classroom and student activities scale were consistent across all items. Secondary school principals reported a decrease in the amount of time spent on these activities. Activities related to classroom enrichment decreased the least, while class trips decreased the most.

The last research question asked: Do secondary school principals' perceptions of the value and impact of state-mandated high-stakes tests on content and mode of instruction differ when examined by school level (e.g., middle and high school) and category (e.g., urban, suburban, and rural)? Significant differences were found when principals' responses were examined by level. Differences were found in the way principals responded to three items: (a) State-mandated high-stakes tests motivate previously unmotivated students, (b) state-mandated high-stakes tests are a fad, and (c) state-mandated high-stakes tests are not an accurate measure of what ESL students know and can do. High school principals' responses to all of the items indicated more agreement to these items except to the statement: state-mandated high-stakes tests are not an accurate measure of what ESL students know and can do. Middle school principals' responses indicated they disagreed more than high school principals did with the item: motivates previously unmotivated students.

A MANOVA was run to explore the differences in secondary school principals' perceptions of the impact of state-mandated tests on content and mode of instruction when examined by school level (e.g., middle/junior high and high school) and category (e.g., urban, suburban, and rural) (see Table 20). Significant differences were found by level in the way principals responded to items on the Tested Areas Scale.

Individual ANOVAs were run to examine how individual items differed by level in the way principals responded to the items on the Tested Areas Scale. Results from the ANOVAs are presented in Table 21. Differences were found in the way principals

responded to questions about (a) areas not covered by state-mandated tests and (b) tested areas without high-stakes attached (see Table 22). The mean scores ranged from $M = 1.930$ to $M = 2.940$. High mean scores of $M > 3.00$ indicate an increase in instructional time, and scores below 3.00 indicate a decrease. High school principals' mean scores were higher, indicating that their schools spent more time on areas not covered by state-mandated tests and tested areas without high stakes. Both middle and high school principals' scores were less than $M = 3.00$, however, indicating that they had a decrease in the amount of time spent in these areas.

Significant differences were found by level in the way principals responded to four items on the Noncore Areas Scale: (a) fine arts, (b) physical education, (c) foreign language, and (d) industrial/vocational education. To explore those differences, the means were examined. A high mean score, $M > 3.00$, indicated that more instructional time was devoted to this area. Overall, the means for middle and high school principals' responses indicated that there had been a slight decrease in the amount of instructional time related to these areas; however, the mean scores for high school principals' responses were slightly higher than the mean scores for the middle school principals' responses in all areas.

When examined by category and level, the principals' responses to the amount of time spent on foreign language were slightly significant. Overall, the responses indicated that there had been a slight decrease in the amount of time spent in foreign language courses. Upon examination, however, it was revealed that in the urban, suburban, and

rural categories, high school principals' responses were higher than middle school principals' responses. This indicated that in a high-stakes testing environment, high schools were spending more time than middle schools were on foreign language courses. Additionally, rural high school principals reported spending more time on foreign language courses than urban or suburban principals did. In contrast, middle school principals reported spending less time on foreign language courses than high school principals did, as a result of state-mandated tests. The mean score for middle school principals in rural schools was higher than the mean scores for urban or suburban middle school principals.

Conclusions

From the findings in this study, the following conclusions were drawn:

1. The principals in this study agreed that state-mandated tests are worthwhile; however, there are negative consequences (e.g., narrowing of curriculum, loss of time due to test preparation, and inaccurate reporting of results) attached to their use. Additionally, principals agreed that media coverage of high-stakes test results does not adequately report how much students are learning and how effective schools are.
2. Principals reported an increase in the amount of time spent in areas that would improve students' performance on state-mandated high-stakes tests. Principals had a strong degree of agreement in reporting a decline in the amount of time

spent on all nontested areas and classroom and student activities. This finding is supported by the results from a recent nationwide survey of teachers (Pedulla et al., 2003).

3. When exploring secondary school principals' perceptions of the value of state-mandated tests by level and category, significant differences surfaced between middle and high school principals' responses to three areas. Both middle and high school principals agreed, however, that state-mandated tests are worthwhile but have negative consequences attached.
4. An examination of secondary school principals' perceptions of the impact of state-mandated high-stakes tests on content and mode of instruction revealed significant differences by level in the amount of time spent on areas not represented in state-mandated high-stakes tests. Middle school principals reported a decrease in the amount of time spent in all areas not covered by state-mandated tests. High school principals reported less of a decrease than middle school principals did in the amount of time spent in these activities.

Discussion

The phenomenon under investigation was a study of middle and high school principals' perceptions of the value and impact of state-mandated high-stakes tests on the content and mode of instruction. Most of the research reviewed for this study involved teachers' and some principals' perceptions of the impact of high-stakes tests on student

learning. The majority of this research was related to teachers. There was very little research related to principals' perceptions of the impact of high-stakes tests.

The secondary school principals who participated in this study indicated that state-mandated high-stakes tests are worth the investment of time and money. The findings in this study are supported by the results from a recent nationwide survey of teachers. The National Board on Educational Testing and Public Policy (Pedulla et al., 2003) conducted a nationwide study to ascertain teachers' attitudes and opinions about state-mandated testing programs. According to the results,

The curriculum standards or frameworks established by states are intended to articulate high expectations for academic achievement and clear outcomes for students. Such curriculum standards have the consequence of establishing homogeneity of course content, thereby focusing classroom instruction and providing teachers with a clear purpose. Regardless of stakes levels, the majority of teachers were positive about their state's content standards or frameworks. Fifty-eight percent of all responding teachers reported that their state-mandated test is based on a curriculum that all teachers should follow. Similarly, more than half of all teachers (55%) reported that if they teach to the state standards or frameworks, students will do well on the state test. (Pedulla et al., 2003, p. 23)

The principals in this study and the teachers who participated in the National Board on Educational Testing and Public Policy study both indicated agreement that accountability systems and high-stakes tests raise academic expectations, but that some unexpected

negative consequences come with them. For example, one of the principals involved in this study wrote on the survey that, “testing is a pain but it has improved instruction!”

It is very clear from the data collected in this study that these principals perceive state-mandated high-stakes tests as having positive consequences, and they also believe that there are some negative consequences attached to them. Some of the negative consequences include (a) a loss of instructional time due to the preparation for high-stakes tests, (b) a narrowed curriculum, and (c) inaccurate reporting of the results by the media.

Based on the findings of this study, it was concluded that middle and high school principals perceive that high-stakes tests have both positive and negative consequences. The results of this study seem to support Stecher’s (2002) argument:

Overall, the evidence suggests that large-scale high-stakes testing has been a relatively potent policy in terms of bringing about changes within schools and classrooms. Many of these changes appear to diminish students’ exposure to curriculum, which undermines the meaning of test scores. It will take more time and more research to determine on balance whether the positive impact on teaching practice and student learning outweigh the negative ones. (p. 100)

The findings from this study are supported by other research studies (Abrams, Pedulla, & Madaus, 2003; Brown, 1993) that reported principals and teachers perceive high-stakes tests as leading to outcomes other than those intended. Most of the unintended consequences of these tests are negative.

Another conclusion of this study was that middle and high school principals perceive that the use of high-stakes tests leads teachers to teach to the test. As a result, areas that are not tested are being given little attention. This conclusion is clearly supported by other studies of teachers and principals indicating that high-stakes tests have led to narrower curricula and less time spent on content that is not covered by the tests (Bolon, 2000; Brown, 1993; Gordon & Reese, 1997; Jinkins, 2001; Mintrop, 2003; Pedulla, 2003; Reed, McDonough, Ross, & Robichaux, 2001). In a more recent study conducted by the National Board on Educational Testing and Public Policy (Pedulla et al., 2003), 60% of the teachers who participated in the study indicated that the time they spent on instruction in tested areas had increased a great deal and the time spent on nontested content had been reduced.

Recommendations for Future Research and Practice

The following recommendations are made based on the findings of this research study:

1. This study should be replicated to include elementary principals to determine if the findings related to principals' perceptions of the value and impact of state-mandated high-stakes tests on instruction may be generalized to all levels.
2. This study should be replicated due to the low reliability level of the tested areas scale. If the replicated study produces the same results, it would validate this study.

3. Research should be done to examine if state-mandated high-stakes tests are having more of a negative impact on elementary and middle schools than on high schools.
4. Further research needs to be done to determine how accountability systems can maximize benefits and minimize negative consequences.
5. Findings from this study need to be made available to policymakers to determine what effects state-mandated tests are having on instruction in Tennessee.
6. Middle school principals and school districts should consider the research related to the benefits of foreign language before reducing the amount of time spent on it. Research shows it is important to offer foreign language instruction as a part of the academic curriculum in the early grades, with continuation through middle and high school.
7. Teachers and principals should not make important decisions about ESL students based on one test. Low scores on standardized tests may mean nothing more than that a learner has not mastered the English language well enough to demonstrate his or her content knowledge and skills on a test.
8. High-stakes decisions should not be made regarding a school or district with high numbers of ESL students based solely on test data. The data may only indicate that a school or district has a high percentage of ESL students and may not be representative of the quality of instruction.
9. Tennessee should revise the manner in which test results are reported to the public. Currently, test results that are published in the newspaper and broadcast on television include just the actual numerical scores. There is very little explanation about the schools,

the populations they serve, or improvements or gains over the previous year. As a result, the parents, the business community, and others tend to use the test scores as a way of comparing the schools to each other. In essence, what they are doing is comparing apples and oranges.

REFERENCES

References

- Abrams, L., Pedulla, J., & Madaus, G. (2003, winter). Views from the classroom: Teachers' opinions of statewide testing programs. *Theory Into Practice, 42*(1), 18-29.
- Acker-Hocevar, M., & Touchton, D. (2001). *Principals' struggle to level the accountability playing field of Florida graded "d" and "f": Schools in high poverty and minority communities*. Paper presented at the annual meeting of the American Educational Research Association, Seattle, WA. (ERIC Document Reproduction Service No. ED458322)
- Adams, L., & Karabenick, S. A. (2000). *Impact of state testing on students and teaching practices: Much pain, no gain?* Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA. (ERIC Document Reproduction Service No. ED443870)
- Ahearn, E. M. (2000). *Educational accountability: A synthesis of the literature and review of a balanced model of accountability*. Alexandria, VA: National Association of State Directors of Special Education. (ERIC Document Reproduction Service No. ED439573)
- Amrein, A. L. (2002). A time-series analysis of the intended and unintended consequences of high-stakes test (Doctoral dissertation, Arizona State University, 2002). *Dissertation Abstracts International, 68*, 1312.

- Amrein, A. L., & Berliner, D. C. (2002). High-stakes testing, uncertainty, and student learning. *Education Policy Analysis Archives*, 10(18). Retrieved July 22, 2003, from <http://epaa.asu.edu/epaa/v10n18/>
- Amrein, A. L., & Berliner, D. C. (2003). The effects of high-stakes testing on student Motivation and learning. *Educational Leadership*, 60(5), 32-38.
- Barksdale-Ladd, M. A., & Thomas, K. F. (2000). What's at stake in high-stakes testing? Teachers and parents speak out. *Journal of Teacher Education*, 51(5), 384-397.
- Bersola, S. H. (2002). The influence of high-stakes standardized tests on school curricula. *Dissertation Abstracts International*, 63(10A), 3427 (UMI No. 3067832)
- Blase, J., & Blase, J. (2004). *Handbook of instructional leadership: How successful principals promote teaching and learning* (2nd ed.). Thousand Oaks, CA: Corwin.
- Bliem, C. L., & Shepard, L. A. (1993). *Parent opinions about standardized tests, teacher's information and performance assessments. A case study of the effects of alternative assessment in instruction, student learning and accountability practices*. Paper presented at the annual meeting of the American Educational Research Association, Atlanta, GA. (ERIC Document Reproduction Service No. ED 378227)
- Bolon, C. (2000). School-based standard testing. *Education Policy Analysis Archives*, 8(23). Retrieved July 22, 2003, from <http://epaa.asu.edu/epaa/v8n23>

- Bossert, S. T., Dwyer, D. C., Rowan, B., & Lee, G. V. (1982). The instructional management role of the principal. *Educational Administration Quarterly*, 18(3), 34-64.
- Bracey, G. W. (2003a). April foolishness: The 20th anniversary of a nation at risk. *Phi Delta Kappan*, 84(8), 616-621.
- Bracey, G. W. (2003b). The 13th Bracey report on the condition of public education. *Phi Delta Kappan*, 85(2), 148-165.
- Brown, D. F. (1993). *The political influence of state testing reform through the eyes of principals and teachers*. Paper presented at the annual meeting of the American Educational Research Association, Atlanta, GA. (ERIC Document Reproduction Service No. ED360737)
- Brown 40 years on. (1994). *The Economist*, 331(7865), 15.
- Brown v. Board of Education of Topeka*. (n.d.). Retrieved March 4, 2004, from <http://11library.thinkquest.org/10718/body.htm>
- Carson, C., Huelskamp, R., & Woodall, T. (1993). Perspectives on education in America. *The Journal of Educational Research*, 86(5), 259-310.
- Ceperley, P., & Reel, K. (1997). The impetus for the Tennessee value-added assessment system. In J. Milman (Ed.), *Grading teachers, grading schools: Is student achievement a valid evaluation measure?* (pp. 163-168). Thousand Oaks, CA: Corwin Press.
- Cimbricz, S. (2002). State-mandated testing and teachers' beliefs and practice.

Educational Policy Analysis, 10(2). Retrieved July 22, 2003, from
<http://epaa.asu.edu/epaa/v10n2>

Clarke, M., Arnold, S., Rhoades, K., Abrams, J. M., & Li, J. (2003). *Perceived effects of state-mandated testing programs on teaching and learning: Findings from interviews with educators in low-, medium-, and high-stakes states*. Boston College: National Board on Educational Testing and Public Policy. (ERIC Document Reproduction Service No. ED456139)

Clarke, M., & Gregory, K. (2003). This issue. *Theory Into Practice*, 42(1), 2-3.

Clarke, M., Haney, W., & Madaus, G. (2000). *High stakes testing and high school completion*. Boston College: National Board on Educational Testing and Public Policy. (ERIC Document Reproduction Service No. ED456139)

Coleman, J., Campbell, E., Hobson, C., McPartland, J., Mood, A., Weinfeld, F., & York, R. (1966). *Equality of educational opportunity*. Washington, DC: U.S. Government Printing Office.

Cooley, V. E., & Shen, J. (2003). School accountability and professional job responsibilities: A perspective from secondary principals. *NASSP Bulletin*, 87(634), 10-25.

Creswell, J. W. (2003). *Research design: Qualitative, quantitative, and mixed methods approaches* (2nd ed.). Thousand Oaks, CA: Sage.

Daniels, H., Bizar, M., & Zemelman, S. (1998). *Best practice: New standards for teaching and learning in America's schools*. Portsmouth, NH: Heinemann.

- Daniels, H., Bizar, M., & Zemelman, S. (2001). *Rethinking high school: Best practice in teaching, learning, and leadership*. Portsmouth, NH: Heinemann.
- Darling-Hammond, L. (1991). The implications of testing policy for quality and equality. *Phi Delta Kappan*, 73(3), 220-225.
- Darling-Hammond, L. (1997). *The right to learn: A blueprint for creating schools that work*. San Francisco: Jossey-Bass.
- Denoyer, R., & White, M. (1992). *Is this really fair: Education reform and ranking districts by test scores*. (ERIC Document Reproduction Service No. ED 350683)
- Dillman, D. A. (2000). *Mail and telephone surveys: The tailored design method*. New York: John Wiley & Sons.
- Dufour, R., & Eaker, R. (1998). *Professional learning communities at work: Best practices for enhancing student achievement*. Bloomington, IN: National Education Service.
- Education Improvement Act, 49 TN 1 SS 601 (1992).
- Elford, G. W. (2002). *Standardized testing: Better accountability for school accountability and management*. Lanham, MD: The Scarecrow Press.
- Evers, W. M., & Walberg, H. J. (Eds.). (2002). *School accountability*. Stanford, CA: Hoover Institution Press.
- Firestone, W. A., Schorr, R. Y., & Monfils, L. F. (Eds.). (2004). *The ambiguity of teaching to the test*. Mahwah, NJ: Lawrence Erlbaum.

- Freund, C. (2002). Sphere of influence: Sputnik's lingering effects. *Reason*, 34, 62.
- Gay, L. R., & Airasian, P. (2003). *Educational research: Competencies for analysis and applications* (7th ed.). Upper Saddle River, NY: Merrill Prentice Hall.
- Gordon, S. P., & Reese, M. (1997). High-stakes testing: Worth the price? *Journal of School Leadership*, 7, 345-368.
- Graglia, L. A. (1996). The triumph and defeat of the nondiscrimination principle. *Society*, 33, 7.
- Graham, G., Wilkens, J. L., Westfall, S., Parker, S., Fraser, R., & Tembo, M. (2002). The effects of high-stakes testing on elementary school art, music, and physical education. *The Journal of Physical Education, Recreation & Dance*, 73(8), 51-54.
- Grant, S. G. (2000). Teachers and tests: Exploring teachers' perceptions of changes in New York state testing program. *Educational Policy Analysis*, 8(14). Retrieved July 22, 2003, from <http://epaa.asu.edu/epaa/v8n14>
- Gulek, C. (2003). Preparing for high-stakes testing. *Theory Into Practice*, 42(1), 42-50.
- Gunzenhauser, M. G. (2003). High-stakes testing and the default philosophy of education. *Theory Into Practice*, 42(1), 51-58.
- Halberstam, D. (2004). Brown v. board of education: What it means to every American. *Parade*, 4-6.
- Hamilton, L. S., Stecher, B. M., & Klein, S. P. (Eds.). (2002). *Making sense of test-based accountability in education*. Santa Monica, CA: RAND.
- Haney, W. (2000). The myth of the Texas miracle in education. *Education*

Policy Analysis Archives, 8(41). Retrieved July 24, 2005, from

<http://epaa.asu.edu/epaa/v8n41>

- Heubert, J. P., & Hauser, R. M. (Eds.). (1999). *High stakes: Testing for tracking, promotion, and graduation*. Washington, DC: National Academy Press.
- Hill, L. C. (2003). Compromise in the classroom. *Vanderbilt Magazine*, 33-39.
- Hoffman, J. V., Assaf, L. C., & Paris, S. G. (2001). High-stakes testing in reading: Today in Texas, tomorrow? *Reading Teacher*, 54(5), 482.
- Holman, L. J. (1995). *The impact of ethnicity, class, and gender on achievement of border area students on a high-stakes examination*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA. (ERIC Document Reproduction Service No. ED384627)
- Holloway, J. H. (2003). Research link: Linking professional development to student learning. *Educational leadership*, 61(3), 85-86.
- Horn, C. (2003). High-stakes testing and students: Stopping or perpetuating a cycle of failure. *Theory Into Practice*, 42, 30-41.
- Jenkins, B. (2001). Collaboration: A road map to success. *Knowledge Quest*, 29(3), 26-27.
- Jones, B. D., & Egly, R. J. (2004). Voices from the frontlines: Teachers' perceptions of high-stakes testing. *Education Policy Analysis Archives*, 8(41). Retrieved July 24, 2005, from <http://epaa.asu.edu/epaa/v8n41>
- Kaplan, L. S., & Owings, W. A. (2001). How principals can help teachers with high-

- stakes testing: One survey's findings with national implications. *NASSP Bulletin*, 85(622), 15-23.
- Kelley, C., Conley, S., & Kimball, S. (2000). Payment for results: Effects of Kentucky and Maryland group-based performance award programs. *Peabody Journal of Education*, 75(4), 159-199.
- Klein, S., Hamilton, L., McCaffrey, D., & Stecher, B. (2000). What do test scores in Texas tell us?. *Education Policy Analysis Archives*, 8(49). Retrieved July 24, 2005, from <http://epaa.asu.edu/epaa/v8n41>
- Kohn, A. (2001). Fighting the tests: A practical guide to rescuing our schools. *Phi Delta Kappan*. Retrieved February 19, 2005, from <http://alfiekohn.org/teaching/ftt.htm>
- Kubow, P. K., & Debard, R. (2000). Teacher perceptions of proficiency testing: A winning Ohio suburban school district expresses itself. *American Secondary Education*, 29(2), 16-25.
- Latimore, R. (2001). The wrath of high-stakes tests. *The Urban Review*, 33(1), 57-67.
- Leedy, P., & Ormod, J. (2001). *Practical research* (7th ed.). Upper Saddle River, NJ: Merrill Prentice Hall.
- Mabry, L., Poole, J., Redmond, L., & Schultz, A. (2003). Local impact of state testing in southern Washington. *Education Policy Analysis Archives*, 11(22). Retrieved July 22, 2003, from <http://epaa.asu.edu/epaa/v11n22>
- Madaus, G. F. (1988). The distortion of teaching and testing: High-stakes testing

and instruction. *Peabody Journal of Education*, 65(3), 29-36.

Madaus, G. F. (1991). The effects of important tests on students. *Phi Delta Kappan*, 73(3), 226-231.

Madaus, G. F., & Clarke, M. (2001). *The adverse impact of high-stakes testing on minority students: Evidence from 100 years of test data*. (ERIC Document Reproduction Service No. ED450183)

Marion, S. F., & Sheinker, A. (1999). *Issues and consequences for state-level minimum competency testing programs*. Minneapolis, MN: National Center on Educational Outcomes. (ERIC Document Reproduction Service No. ED440515)

Maurice, J. M., & Karr-Kidwell, P. J. (2003). *Accountability issues for instructional leadership: Field research, site-based management, and a campus action plan*. (ERIC Document Reproduction Service No. ED473813)

McNeil, L. M. (2000). *Contradictions of School Reform: Educational Costs of Standardized Testing*. New York: Routledge.

Mehrens, W. A., & Lehmann, I. J. (1987). *Standardized tests in education* (4th ed.). New York: Longman.

Merriam, S. B. (1998). *Qualitative research and case study applications in education*. San Francisco: Jossey-Bass.

Mintrop, H. (2003). The limits of sanctions in low-performing schools: A study

of Maryland and Kentucky schools on probation. *Education Policy Analysis Archives*, 11(3). Retrieved July 22, 2003, from <http://epaa.asu.edu/epaa/v11n3.html>

Mitchell, K. (1996). *Reforming and conforming: NASDC principals discuss school accountability systems*. Retrieved July 22, 2003, from <http://www.rand.org/publications/mr/mr716/>

Morrison, L. (1991). *Why conventional education reform fails: The case for market-based restructuring*. Harrisburg, PA: Commonwealth Foundation for Public Policy. (ERIC Document Reproduction Service No. ED397210)

Mulvenon, S. W., Connors, J. V., & Lenares, D. (2001). *Impact of accountability and school testing on students: Is there evidence of anxiety?* Paper presented at the annual meeting of the Mid-South Educational Research Association, Little Rock, AR. (ERIC Document Reproduction Service No. ED460155).

Nardi, P. M. (2006). *Doing survey research: A guide to quantitative methods*. Boston: Pearson Education.

National Commission on Excellence in Education. (1983, April). *A nation at risk: The imperative for educational reform*. Washington, DC: U.S. Government Printing Office.

Orlich, D. (2003). An examination of the longitudinal effect of the Washington

Assessment of Student Learning (WASL) on student achievement. *Educational Policy Analysis*, 11(8). Retrieved Sept. 14, 2005, from <http://epaa.asu.edu/epaa/v11n18>.

Parker, J. B. (1994). *Tennessee teachers and testing: A profile of perceptions*.

Unpublished doctoral dissertation, The University of Tennessee, Knoxville.

Patterson, F. E. (2002). *Educational values and accountability in Tennessee: Ethical dilemmas and moral imperatives*. Unpublished doctoral dissertation, The University of Tennessee, Knoxville.

Pedulla, J. (2003). State-mandated testing: What do teachers think? *Educational Leadership*, 61(3), 42-46.

Pedulla, J., Abrams, L., Madaus, G., Russell, M., Ramos, M., & Miao, J. (2003).

Perceived effects of state-mandated testing programs on teaching and learning:

Findings from a national survey of teachers. Chestnut Hill, MA: Boston College,

Center for the Study of Testing, Evaluation, and Educational Policy.

Perlman, C. L. (1985). *Results of a citywide testing program audit in Chicago*.

Paper presented at the annual meeting of the American Educational Research

Association, Chicago, IL. (ERIC Document Reproduction Service No.

ED263212)

Popham, J. (2001). Teaching to the test? *Educational Leadership*, 58(6), 16-20.

Rapp, D. (2002). National board certification teachers in Ohio give state

education policy, classroom climate, and high-stakes testing a grade of

f. *Phi Delta Kappan*, 84(3), 215-218.

Reed, C. J., McDonough, J., Ross, M., & Robichaux, R. (2001). *Principals' perceptions of the impact of high-stakes testing on empowerment*. Paper presented at the annual meeting of the American Educational Research Association, Seattle, WA. (ERIC Document Reproduction Service No. ED459538)

Rentner, D. S., Chudowsky, N., Fagan, T., Gaylor, K., Hamilton, M., & Kober, N. (2003). *From the capital to the classroom: State and federal efforts to implement the No Child Left Behind Act*. Washington, DC: Center on Educational Policy.

Rentner, D. S., & Hamilton, M. (2003). First signs of the new accountability. *Principal Leadership*, 3(9), 10-12.

Sanders, W. L., & Horn, S. P. (1998). Research findings from the Tennessee value-added assessment system (TVAAS) database: Implications for educational evaluation and research. *Journal of Personnel Evaluation in Education*, 12(3), 247-256.

Schleisman, J. (1999). *An in-depth investigation of one school district's responses to an externally mandated, high-stakes testing program in Minnesota*. Paper presented at the annual meeting of the University Council for Educational Administration, Minneapolis, MN. (ERIC Document Reproduction Service No. ED440465)

Shepard, L. A. (1991). Will national tests improve student learning? *Phi Delta Kappan*, 73(3), 232-238.

Shepard, L. A., & Cutts-Dougherty, K. C. (1991). *Effects of high-stakes testing on*

instruction. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL. (ERIC Document Reproduction Service No. ED337468)

Sidney, H. (2003). Oct. 4, 1957: The space race lifts off. *Time*, 161, A32.

Singh, J., & McMillan, J. H. (2002). *Staff development practices in schools demonstrating significant improvement on high-stakes tests*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA. (ERIC Document Reproduction Service No. ED464103)

Smith, M. J., Detch, E. R., & Morgan, J. G. (2004). *The education improvement act: A progress report*. Nashville, TN: State of Tennessee Comptroller of the Treasury.

Snow-Renner, R. (2001). *Teachers' perspectives on standards-based education: Initial findings from a high-performing, high-needs, school districts*. Aurora, CO: Mid-continent Research for Education and Learning. (ERIC Document Reproduction Service No. ED457150)

Stake, R. E. (1991). The teacher, standardized testing, and prospects of revolution. *Phi Delta Kappan*, 73(3), 243-247.

Stecher, B. M. (2002). Consequences of large-scale, high-stakes testing on school and classroom practices. In L. S. Hamilton, B. M. Stecher, & S. P. Klein (Eds.), *Making sense of test-based accountability in education* (pp. 79-100). Santa Monica, CA: RAND.

Stecher, B. M., & Barron, S. I. (1999). *Quadrennial milepost: Accountability testing in*

Kentucky (CSE Tech. Rep. No. 505). Los Angeles, CA: National Center for Research on Evaluation, Standards, and Student Testing.

Stecher, B. M., & Barron, S. I. (2003). *Unintended consequences of test-based accountability when testing in "milepost" grades*. Santa Monica, CA: RAND.

Stedman, J. B., Irwin, P. M., Lyke, B., & Riddle, W. C. (1990). *National education goals: Where are we now?* Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA. (ERIC Document Reproduction Service No. ED359634)

Stronge, J. H., & Tucker, P. D. (2000). *Teacher evaluation and student achievement. Student assessment series*. Washington, DC: National Education Association.

Taylor, K. L. (2003). Through the eyes of students. *Educational Leadership*, 60(4), 72-75.

Tennessee Department of Education. (2003a). *Executive summary: Accountability workbook*. Nashville, TN: Author.

Tennessee Department of Education. (2003b). *No child left behind: A handbook for principals*. Nashville, TN: Author.

Tennessee Department of Education. (2005). *Tennessee-sde directory- school list*. Retrieved October 20, 2005, from <http://www.k-12.state.tn.us/sde/CreateSchoolList.asp?status=A&schlevel=000,003>

Tennessee Small School Systems v. McWherter, 821 S.W. 2nd 139 (Tenn. 1993). (ERIC Document Reproduction Service No. ED460076)

- Turchi, L., Johnson, D., Owens, D. D., & Montgomery, D. (2002). *The impact of accountability on the professional development of teachers: Preliminary evidence from case studies in six southern states*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA. (ERIC Document Reproduction Service No. ED469948)
- U.S. Department of Education. (2003). *NAEP 2003 mathematics assessment*. Jessup, MD: Author.
- U.S. Department of Education. (2004). *No child left behind: A toolkit for teachers*. Jessup, MD: Author.
- Vogler, K. (2002). The impact of high-stakes, state-mandated student performance assessment on teachers' instructional practices. *Education*, 123(1), 39-55.
- Walker, S. F. (2000). *High-stakes testing: Too much? Too soon?* Denver, CO: Education Commission of the States. (ERIC Document Reproduction Service No. ED45219)
- Wiggins, R. A. (1994). *The principal as instructional leader: Inducement or deterrent to teachers' personal professional growth?* Paper presented at the annual meeting of the Association for Supervision and Curriculum Development, Chicago, IL. (ERIC Document Reproduction Service No. ED380886)
- Williams, K. L. (2002). Accountability and professional prerogative: The impact of high stakes testing on teacher agency (Doctoral dissertation, California State University, 2002). *Dissertation Abstracts International*, 63, 08A.

Williams, N. M. (2003). Thinking outside the bubble. *Educational Leadership*, 61(3), 82-83.

Wilson, B., & Corbett, H. D. (1991). *Two state minimum competency testing programs and their effects on curriculum and instruction*. Philadelphia, PA: Research for Better Schools. (ERIC Document Reproduction Service No. ED377251)

Young, T. A. (1996). A study of the perceptions of Tennessee teachers, principals, superintendents, legislators, and Department of Education representatives regarding the impact of the Tennessee value-added assessment system (Doctoral dissertation, University of Alabama, 1996). *Dissertation Abstracts International*, 57, 2317.

Zellner, L., & Jinkins, D. (2001). *Consequences of high-stakes testing on the literacy programs of high-performing learning communities*. Paper presented at the 51st National Reading Conference. (ERIC Document Reproduction Service No. ED465810)

APPENDICES

Appendix A

Survey Modification Form

IMPORTANT!!!!

Survey Modification Form

Please take this opportunity to assist me in making any needed modifications to the survey instrument by addressing the items below.

1. Do you think the questions in this survey adequately represent the topic?

Yes

No

If no, please explain. _____

2. Please provide any comments that would improve the layout or clarity of the questions or instructions in the survey instrument.

Comments:

Thank you for your time!

Appendix B

Prenotification Letter

April 11, 2006

Dear Principal:

A few days from now you will receive in the mail a request to fill out a brief survey for an important statewide study. The purpose of the study is to explore secondary school principals' perceptions of the impact of state-mandated, high-stakes tests on instruction in Tennessee.

I am writing in advance because we have found that many people like to know ahead of time that they will be contacted. This study is an important one that will help policymakers, educators, the public, and others understand the perceived impact of state-mandated, high-stakes tests on instruction in Tennessee's public secondary schools. The results of the study will be made available to policymakers for use in framing accountability policy for the state.

Thank you for your time and consideration. It is only with the generous help of principals like you that this research can be successful.

Sincerely,

Clifford Davis, Jr.
Principal

Appendix C

Sample Letter to School Principals

April 2006

Dear Principal:

I am an administrator with the Knox County School System and a doctoral student working with Dr. E. Grady Bogue at the University of Tennessee, Knoxville. As a part of my dissertation, I am conducting a survey. The purpose of this research is to determine secondary school principals' perceptions of the value and impact of high-stakes tests on content and mode of instruction in Tennessee.

Your participation in this study consists of completion of a survey concerning the impact of state-mandated high-stakes tests on instruction. Completion of this survey should require about 15 minutes of your time. There is minimal risk involved in participating in this study. Your identity will be kept confidential. The survey instrument does have an identification number. This will enable me to determine which surveys have not been returned. No individual responses will be revealed, and all data will be reported in a composite form.

If you have questions at any time about the study or the procedures, you may contact the researcher, Clifford Davis, Jr., at Kams High School, (865) 539-8670, ext. 163, or by e-mail (davic29@k12tn.net). If you have questions about your rights as a participant, contact the compliance section of the Office of Research at the University of Tennessee Knoxville at (865) 974-3466.

Please complete and return the enclosed questionnaire using the stamped addressed envelope by April 25, 2006. Thank you for your cooperation.

Sincerely,

Clifford Davis, Jr.
Knox County Schools

Appendix D

Permission to Use Survey Instrument

Dear Mr. Davis,

Feel free to use the Teacher's Survey. The only request that the National Board has is to see your completed work and your results. It would be interesting to see your results. Good luck with your dissertation, please let us know if there is anything else we can help you with, Cindy Yang.

| On Wed, 6 Jul 2005 22:40:06 -0400

Appendix E

Survey Instrument

Instructions for completing the survey: The purpose of this study is to explore secondary school principals' perceptions of the value and impact of state-mandated high-stakes tests on content and mode of instruction in Tennessee. These issues are particularly important for the education reform efforts that are currently taking place.

Your individual responses will be kept strictly confidential and will not be provided to any other person or group. Since you have been selected as part of this study, your responses are very important.

1. Which category best describes your school?

- Urban
- Suburban
- Rural

2. How do your school's results on the state-mandated test compare to those of other schools in your state?

- Above average
- Average
- Below Average

3. Which level best describes your school?

- Middle School/Junior High
- High School

4. In which part of the state is your school located?

- East
- Middle
- West

5. Which category best describes the size of your school?

- 1 - 499
- 500 - 999
- 1000 - 1499
- 1500 - 1999
- 2000 +

Please indicate the extent to which you agree with each of the following statements (#6-#18) by filling in the box that corresponds with your response.

	Strongly Agree	Agree	Disagree	Strongly Disagree
6. Overall, the benefits of the state-mandated testing program are worth the investment of time and money.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7. Media coverage of state-mandated testing issues accurately reflects the quality of education in the state.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
8. Scores on the state-mandated test accurately reflect the quality of education students have received.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9. The state-mandated test has brought much needed attention to education issues in my district.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
10. The state-mandated test is as accurate a measure of student achievement as a teacher's judgment.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
11. The state-mandated test motivates previously unmotivated students to learn.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
12. The state-mandated test measures high standards of achievement.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
13. The state-mandated testing program is just another fad.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
14. Media coverage of state-mandated testing issues has been unfair to teachers.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
15. Media coverage of state-mandated testing issues adequately reflects the complexity of teaching.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
16. Teachers in my school have found ways to raise state-mandated test scores without really improving student learning.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
17. The state-mandated test is NOT an accurate measure of what students who are acquiring English as a second language know and can do.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
18. Score differences from year to year on the state-mandated test reflect changes in the characteristics of students rather than changes in school effectiveness.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

19. In what ways, if any, has the amount of time spent on each of the following activities changed in your school in order to prepare students for the state-mandated testing program?

	Decreased a Great Deal	Moderately Decreased	Stayed About the Same	Moderately Increased	Increased a Great Deal
Instruction in tested areas	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Instruction in areas not covered by the state-mandated test	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Instruction in tested areas with high stakes attached (e.g., promotion, graduation, teacher rewards)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Instruction in tested areas without high stakes attached	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Instruction in the fine arts	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Instruction in physical education	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Instruction in foreign language	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Instruction in industrial/vocational education	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Student free time (e.g., recess, lunch)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Field trips (e.g., museum tour, hospital tour)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Class trips (e.g., circus, amusement park)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Student choice time (e.g., games, computer work)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Organized play (e.g., games with other classes)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Enrichment school assemblies (e.g., professional choral group performances)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Administrative school assemblies (e.g., awards ceremonies)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Classroom enrichment activities (e.g., guest speakers)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Student performance (e.g., class plays)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Parental contact	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Background Information

20. How many years of administrative experience do you have including this year?

- 1 4-8 13-20
 2-3 9-12 Over 20

21. What is your gender? Female Male

22. Please mark the appropriate range of your age.

- 20-30 41-50 60+
 31-40 51-60

23. Mark ALL of the following categories that best describe you.

- African American
 American Indian or Alaskan Native
 Asian
 White
 Pacific Islander
 Hispanic
 Other, please specify: _____
-

Thank you for your cooperation with this study!

Your feedback is greatly appreciated!

Appendix F

Characteristics of Survey Respondents

Respondents Characteristics		N	% of Respondents
Gender	Male	77	22.5
	Female	221	73.2
Age	20-30	1	.3
	31-40	43	14.2
	41-50	70	23.2
	51-60	157	52.0
	60+	27	8.9
Race/Ethnicity	African American	34	11.4
	American Indian/ Native American	1	.3
	White	263	88.6
	Asian/Pacific Islander	0	.0
	Hispanic	1	.3
School Level	Middle/Jr. High School	138	45.7
	High School	161	53.3
School Category	Urban	58	19.2
	Suburban	79	26.4
	Rural	162	53.6
School Performance	Above average	121	40.1
	Average	156	51.7
	Below average	18	2.3
School Size	1-499	92	30.5
	500-999	134	44.4
	1000-1499	46	15.2
	1500-1999	19	6.3
	2000+	8	2.6
Years of Administrative Experience	1	4	1.3
	2-3	20	6.6
	4-8	93	30.8
	9-12	55	18.2
	13-20	66	21.9
	Over 20	60	19.9

Appendix G

Thank You/Reminder Postcard

May 1, 2005

Two weeks ago a survey seeking your opinions about the impact of state-mandated, high-stakes tests on content and mode of instruction in Tennessee was mailed to you. Your name was taken from a list of all middle and high school principals in the state of Tennessee.

If you have already completed and returned the survey to me, please accept my sincere thanks. If not, please do so today. I am especially grateful for your help because it is only by asking educators like you to share your opinions that we can truly understand how state-mandated policies are impacting teaching and learning.

If you did not receive a survey, or if it was misplaced, please call me at (865)539-8670 (extension 163) and I will send you another via U.S. mail today.

Clifford Davis, Jr.
Principal of Karns High School
2710 Byington-Solway Road
Knoxville, TN 37931

Appendix H

Second Sample Letter to Principals

May 25, 2006

About four weeks ago I sent you a survey that asked about your opinion of the impact of high-stakes on content and mode of instruction in Tennessee. To the best of my knowledge, it has not been returned.

The response rate from principals involved in the study has been great. I believe that the results will be useful to policymakers, educators, and others.

I am writing again because of the importance your response is in helping me get accurate results. It is only with the participation of the majority of the principals in public schools in the state that I can be sure that the results are representative.

A survey identification number is printed on the back of the survey so I can check your name off as the surveys are returned. After the student has been completed the list of names of participants will be destroyed so that so that individual names can never be connected to the results in any way. Protecting the confidentiality of the participants in the study is very important to me.

I hope that you will complete and return the survey soon. If for any reason, you choose not to complete it, please let me know by returning a note or blank survey in the enclosed stamped envelope.

Sincerely,

Clifford Davis, Jr.

P.S. If you have any questions, please feel free to call me at (865) 539-8670, extension 163.

Appendix I

Content of Instruction

Content Items	Scale
Tested areas	Tested Areas
Tested areas not covered by state mandated tests	Tested Areas
Tested areas with high stakes	Tested Areas
Fine arts	Non-core
Foreign language	Non-core
Industrial/vocational	Non-core
Physical education	Non-core

Appendix J

Mode of Instruction

Mode of Instruction
Basic Skills
Classroom enrichment
Concept development
Cooperative learning
Critical thinking
Field trips
Individual seat work
Using problems similar to those on the test
Whole group instruction

VITA

Clifford Davis, Jr. was born in Myrtlewood, Alabama on July 5, 1965. He graduated in 1983 from Linden High School in Linden, Alabama. In 1987, he graduated from Alabama A.&M. University in Huntsville, Alabama with a Bachelor of Science in mathematics education. In 1993, he earned his masters of mathematics degree from the University of Tennessee, Knoxville. In 1995, he earned his Educational Specialist degree in administration and supervision from the University of Tennessee, Knoxville. He also earned his doctoral degree in Educational Administration and Supervision from the University of Tennessee, Knoxville, in December 2006.

Most of his work experience has been in secondary education. He has served as a high school mathematics teacher, middle school mathematics teacher, mathematics consultant, middle school assistant principal and middle school principal. He currently serves as principal of Karns High School in Knoxville, Tennessee.