

AN ANALYSIS OF ELEMENTARY TEACHERS'
ATTITUDES TOWARD AND THE INSTRUCTIONAL
IMPACT OF STANDARDIZED TESTS
THESIS

Submitted to

The School of Education and Allied Professions of the

UNIVERSITY OF DAYTON

In Partial Fulfillment Of
The Requirements for the Degree

Master of Science in Education

By

Joseph K. Ballengee

UNIVERSITY OF DAYTON

Dayton, Ohio

November 2000

APPROVED BY:

Official Advisor

Faculty Reader

Faculty Reader

ABSTRACT

AN ANALYSIS OF ELEMENTARY TEACHERS' ATTITUDES TOWARD AND THE INSTRUCTIONAL IMPACT OF STANDARDIZED TESTS

Name: Ballengee, Joseph K.
University of Dayton, 2000

Advisor: Patricia Grogan, Ph.D.

This thesis is a quantitative and descriptive qualitative study of the attitudes of elementary teachers toward standardized tests and the relationship between the expressed attitudes and the impact of standardized tests on instructional decisions. The study was conducted in one suburban elementary school in Virginia. The participants in this study included 17 teachers of grades two through six, two teachers of learning disabled students, and one art teacher. All 20 teachers completed questionnaire surveys that assessed their attitudes towards standardized tests. Teachers of grades three, four, and five (12 in all) were interviewed by the researcher to determine the impact standardized testing has had on their instructional decisions. All of the teachers interviewed have made adjustments in their classroom instruction to accommodate the standardized tests. The analysis indicated a relationship between the attitude expressed and the amount of time spent preparing the students for the tests. In general, the more negative the attitude, the more time was spent on test preparation. Conversely, the more positive the attitude, the less time was spent on test preparation.

ACKNOWLEDGEMENTS

I would like to thank my wife, Kelli, for her loving support, patience, and proofreading skills during my work on this research project. I would also like to thank my children, Cicely and Mason, for their understanding and support.

A special thanks to Steve and Lisa Nordmeyer and family of Beavercreek, Ohio, for opening their home to me during the preparation stages of my research proposal.

I thank my advisor, Dr. Patricia Grogan, for her guidance and willingness to work with me "long distance" so that I could complete the research at home in Virginia.

Finally, I want to thank the faculty of E. Barbour Hutchison Elementary School in Herndon, Virginia for their willing participation in my research and especially the principal, Frank Mehm, for allowing me the opportunity to work within his school.

TABLE OF CONTENTS

ABSTRACT.....	iii
ACKNOWLEDGEMENTS.....	iv
LIST OF TABLES.....	vii
LIST OF FIGURES.....	vii
CHAPTER:	
I. INTRODUCTION TO THE PROBLEM.....	1
Purpose for the Study.....	1
Problem Statement.....	2
Hypothesis.....	2
Assumptions of the Study.....	3
Delimitations of the Study.....	3
Limitations of the Study.....	3
Definition of Terms.....	4
Summary.....	4
II. REVIEW OF THE LITERATURE.....	6
History of Standardized Tests.....	6
Uses of Standardized Tests.....	7
Problems Associated with Standardized Tests.....	10
Reliability Concerns.....	12
Validity Concerns.....	13
Teacher Attitudes toward Standardized Tests.....	14
Impacts of Standardized Testing on Instruction.....	21
Summary.....	23
III. METHODS.....	25
Subjects.....	25
Setting.....	25
Data Collection.....	26
Treatment.....	27
Summary.....	28
IV. RESULTS.....	29
Presentation of Survey Results.....	29
Presentation of Interview Results.....	32
Summary.....	35

V.	DISCUSSION	36
	Survey Results.....	36
	Impact on Teaching and Instructional Decisions	44
	Relationship Between Attitude and Impact on Instructional Decisions.....	48
	Summary	51
	Conclusions.....	52
	APPENDIX A: Standardized Test Attitude Questionnaire.....	54
	APPENDIX B: Interview Guide	57
	REFERENCES	58

LIST OF TABLES

1. Subject Response by Grade Level.....	28
2. Standardized Test Attitudes of Teachers Expressed as Percentile Scores.....	30
3. Time Spent on Test Preparation/Review.....	33
4. Assessment of Attitude.....	41
5. Analysis of Attitude by Teacher.....	42
6. Teacher Attitude by Grade.....	43
7. Activities Eliminated or Reduced.....	45
8. Impact on Instructional Decisions.....	47
9. Comparison Between Time Spent on Test Preparation and Attitude Scores.....	49

LIST OF FIGURES

1. Attitude by Years Taught.....	43
2. Percent of Time Spent Compared to Negative Responses.....	50
3. Percent of Time Spent Compared to Positive Responses.....	50

CHAPTER I

INTRODUCTION TO THE PROBLEM

Purpose for the Study

Standardized tests are being used in schools across the country at an increasing rate. Legislators, parents, and others are seeking to reform the schools, and standardized tests are seen as a convenient method to measure how students are performing. Haney and Madaus (1989) quote Chris Pipho of the Education Commission of the States as saying "Nearly every large education reform effort of the last few years has either mandated a new form of testing or expanded uses of existing testing" (p.684).

But standardized tests are not new. They have been used in the United States for over 100 years. (Haney & Madaus, 1989). Now, however, the test scores are being used for purposes other than originally intended. In some states, test scores are used as a basis for firing teachers and/or administrators, for school funding decisions, for rewarding or punishing schools and for other top-level decisions regarding schools and those who work in them. In addition, legislators, parents, and educators use standardized test scores as a way to compare student performance among schools and among classrooms.

Test-based comparisons of schools, districts, and states have increased the pressure on educators to make sure their students perform well on standardized tests. In the defense of standardized tests, some argue the tests measure objectives that are important for students to know. Critics argue that relying on test results distorts education. They contend that important objectives are being ignored in the classroom because they are not included on tests that count (Gronlund & Linn, 1990).

Teachers are caught up in the reform movement, too. While they generally support reform, they have reservations about the use of standardized tests (Stake, 1991). They see the tests as but one indicator of student achievement. They believe that they, themselves, are the best source of evaluation because they see the students on a day-to-day basis and observe them in a variety of situations and learning environments (Salmon-Cox, 1981). Standardized tests only measure one aspect of student achievement.

The use of standardized tests has had a definite impact in the classroom. Many teachers are altering their instructional styles and content. "Teaching to the test," modifying lessons to emphasize testable material, and even dropping subject matter not covered by tests are examples of the kinds of impact reported in research studies. Some teachers feel forced to make instructional decisions based on their perceptions of the importance of standardized test scores. Since teacher attitudes and instructional decisions may be influenced by the use of standardized tests, it becomes important to identify how much of an impact standardized tests have in the classroom.

Problem Statement

The purpose of this study, then, was to determine the relationship, if any, between the attitudes of teachers, in a selected elementary school, towards state-mandated standardized tests and how the standardized tests affect the teachers' instructional decisions.

Hypothesis

It was hypothesized that there would be a positive correlation between the degree of negative attitude toward standardized tests and the degree to which teachers' instructional decisions are affected.

Assumptions of the Study

The following assumptions were made for the conduct of this study.

1. The survey questions were valid.
2. The subjects answered the survey questions honestly.
3. The subjects were honest during the interview.
4. The subjects were interested and willing to complete the survey and participate in the interview.
5. The qualitative methods used to analyze the data yielded an accurate correlation between the data.
6. The subjects were qualified elementary teachers.

Delimitations of the Study

The delimitations of the study were as follows.

1. The subjects were teachers of grades two through six in one suburban Virginia public elementary school.
2. The study was limited to one survey for each teacher and one interview with each third, fourth, and fifth grade teacher.
3. The survey and the interview were the only data-gathering methods used.

Limitations of the Study

The limitations of the study were as follows.

1. The findings may apply only to the selected sample subjects.
2. The findings were dependent upon the subjects completely answering the survey.

3. The findings were dependent upon the subjects answering the survey and interview honestly.

4. The findings may reflect emotional responses to a recently implemented high-stakes standardized test within the Virginia school system.

Definition of Terms

1. Attitude is the state of mind, behavior, or conduct, as indicating one's feelings, opinions, or purpose (Arnott, 1992).
2. Elementary teachers are teachers who have a bachelors degree (or higher) in elementary education (grades one through six or one through eight) and are currently employed in teaching positions (Arnott, 1992).
3. Likert Scale is a self-report method for measuring attitude. It lists clearly favorable or unfavorable statements. The subjects respond to each statement on the following five-point scale: strongly agree (SA), agree (A), undecided (U), disagree (D), and strongly disagree (SD) (Arnott, 1992).
4. Instructional decisions are the day-to-day decisions a teacher makes concerning what subject matter, curriculum, test preparation, and methodology will be used in the classroom to meet learning objectives.

Summary

In this chapter, the purpose of the study, problem statement, hypothesis, and assumptions were stated. Also, the delimitations and limitations of the study were set forth. Finally, the chapter defined the terms used throughout the study. In the following chapter, the review of literature on this subject is examined to show previous work done

to assess teacher attitudes towards standardized tests and the impact standardized tests have had in classrooms in other high-stakes test environments.

CHAPTER II

REVIEW OF THE LITERATURE

History of Standardized Tests

Although the use of standardized testing has been increasing in recent years, they are nothing new. In fact, standardized tests have been used in America for over 100 years. Before the turn of the century, most standardized tests were criterion-referenced tests. These tests measured basic skills in instructional areas. Norm-referenced tests saw increasing use in the 1920's and 1930's, partially as an outgrowth of Army job testing done in World War I (Haney & Madaus, 1989).

Historically, the use of standardized tests dates back to at least 1444. They were used in Italy to help determine teacher salaries. The salaries were partly dependent on how a teacher's pupils performed on tests related to specific areas of the curriculum. Later, teacher pay in England and the English colonies was related to student performance. By 1920, however, these programs were essentially abandoned. There was a brief revival of this practice in the United States in the 1960's, but it did not last long (Madaus & Stufflebeam, 1984).

After World War II, school districts used standardized achievement tests to make inferences about program effectiveness, to identify specific program weaknesses, and to evaluate school curriculum and the overall performance of the educational system (Madaus & Stufflebeam, 1984).

There was another increase in the use of standardized testing in the 1950's and 1960's. This was directly tied to Federal legislation that promoted widespread testing. The

National Defense Education Act of 1958 and the Secondary Education Act of 1965 were primarily responsible for this upsurge (Haney & Madaus, 1989).

Then, in the 1970's there was a movement to demand accountability in schools, again increasing the use of standardized tests. The impetus was the fact that there was a lot of money being spent on schools, but the Scholastic Aptitude Test scores were dropping. People demanded that schools prove the money was well spent (Brandt, 1989).

There is empirical evidence for the increasing use of standardized tests. Resnick (1981) reports that test sales in constant dollars have continued to rise in the period since World War II and have climbed faster than student enrollments. In fact, according to Sacks (1997), sales of standardized tests to public schools more than doubled between 1960 and 1987, while enrollments were up just 15%. He estimates that as many as 127 million tests per year are administered at the K-12 levels alone.

Urduan and Paris (1994) cite a 1993 report by Madaus and Tan which traces the growth in achievement testing to federal legislation and state-mandated testing programs that are intended to promote higher educational standards and better systems of accountability. They point out "...the stakes of standardized test results have increased as well. State and federal funding decisions, teacher salaries and promotion, and school programs often hinge on standardized test results" (p. 138).

Uses of Standardized Tests

Sproull and Zubrow (1981) provide three categories of standardized tests: 1) achievement tests, 2) ability tests, and 3) vocational aptitude tests. Achievement and ability tests are used at almost the same frequency, while vocational aptitude tests are

used much less frequently. All school systems employ at least one achievement test; parochial and private school systems use achievement tests more than any other kind.

According to Sproull and Zubrow (1981), there are four justifications commonly given for the use of standardized tests. The first is for the diagnosis and placement of individual students. The second is for evaluating instructional programs. Thirdly, tests are justified for use as a measure of student end-of-year achievement. Finally, test results are reported to outside audiences (parents, legislatures, etc.).

Jaeger (1991) provides additional rationale. State legislatures have set educational reform goals. These posit that all students need certain basic skills to be productive citizens and must possess a body of certain knowledge and skills. Students also need to prove mastery of specific skills so that they will be able to function in the workplace. The mandated competency tests reflect these goals.

The use of standardized tests has expanded beyond these basic goals. In Texas, schools are ranked based on their students' standardized test scores. "Improvement" in schools is measured by comparing consecutive test scores. The state may intervene in a school if it shows consistently low scores. On the other hand, the state may reward schools that have high scores or show improvement in scores over time. Students are also tracked into remedial classes based on their individual test scores (Gordon & Reese, 1997). Likewise, Indiana bases school accreditation and provides incentive rewards (including funding and recognition) based on school performance on standardized tests. Like Texas, students are evaluated for summer remediation based on their individual scores (Gilman & Reynolds, 1991).

North Carolina uses standardized tests to define the purposes of the state's program of education. It requires students to pass a competency test to qualify for graduation from high school. In general, standardized tests play a part in the state's efforts toward school accountability (Jaeger, 1991).

Similarly, Alaska uses standardized tests to measure the performance of the state's educational system. It maintains a database to aid education policy makers. It also uses scores to determine the number of students who need remedial assistance with basic skills (Jaeger, 1991).

States and school systems have evolved additional uses for standardized tests. Some states and teachers mold their curriculum to reflect what is covered on the mandated tests. Some teachers focus their instruction on material on the tests. Test results are used to group or track students based on their scores. Finally, teachers use test scores to supplement information they already possess about their individual students (Salmon-Cox, 1981).

The various uses of standardized tests have led to some problems. In Kentucky, tests were originally used to identify students who needed additional help. But, when newspapers published the test scores and ranked the schools based on their scores, parents began putting pressure on superintendents to improve the schools. The superintendents put pressure on the principals, who put pressure on the teachers. As Madaus said, "The point is that the tests were being used in a way different from the original purpose and were perceived as accountability tools. This changed the nature of the testing program" (Brandt, 1989, p.27).

Problems Associated with Standardized Tests

Standardized tests have been sold as scientifically developed instruments that objectively, simply, and reliably measure students' achievements, abilities, or skills. However, some question the assumptions behind the construction and use of standardized tests. Neill and Medina (1989) assert that the studies conducted to determine the reliability and validity of standardized tests are inadequate. They maintain that tests are administered in environments that erase claims of standardization. By this, they mean that the test preparation provided by a teacher, or any deviation from strict administrative procedures, introduce variables into the scores that invalidate the standardization aspect. In addition, they feel standardized tests discriminate against minority students, or students from low-income families.

Neill and Medina (1989) also feel that the writers of the tests assume that student knowledge, skills, and abilities can be isolated, sorted to fit on a linear scale, and reported as a single score. In addition, they feel the writers assume that the knowledge, skills, and abilities tend to be distributed on a normal bell curve, which they maintain is not the case. They maintain these assumptions are erroneous and therefore erode the validity of the tests.

In general, critics charge that standardized tests give false information about the status of learning in the nation's schools. Standardized tests are seen as unfair to (or biased against) some kinds of students, namely, minority, low-income, and those with limited English. Standardized tests tend to corrupt the process of teaching and learning. And, standardized tests focus time, energy, and attention on simpler skills that are easily

tested and away from higher-order thinking skills and creative endeavors (Haney & Madaus, 1989).

Over the years, studies show that teachers do not find standardized achievement test scores very useful. Teachers report that standardized tests seldom report talents that they had not already recognized. They also feel standardized tests seldom provide them with diagnostic information that helps them redirect their teaching (Stake, 1991).

Resnick (1981) found standardized tests were not relevant to teachers because they only measure certain aspects of the teachers' cognitive goals for students and none of the social goals. Also, the tests measure to some extent students' current achievement or level of skill, but do not provide the broad-based measurement (problem solving, creative thinking, social interaction, etc.) that teachers prefer.

Smith's (1991a) following comments summarize the problems with standardized tests:

Standardized achievement tests are designed as measures of individual pupil progress in relation to national peers. If external audiences use [standardized] test scores as measures of school effectiveness and accountability and as triggers for reform, school personnel will focus their efforts on improving the scores without respect to, and to the neglect of, other equally plausible and valuable outcomes. The boosted indicator will not likely generalize to alternative indicators, such as the number and quality of books the children read, their writing, projects they undertake, or even to other achievement tests. When an indicator is so fallible that it changes in relation to short-term test preparation and test-wiseness training or the social and ethnic composition of the population, it is worth little in public

debate over school effectiveness or in the disbursing of rewards and punishments from society (p. 541).

Reliability Concerns

Neill and Medina (1989) observe that test writers rarely conduct reliability tests over time to measure the consistency of standardized tests. They also believe nonstandard procedures for administering tests and the impact examiners may have on students undermine test reliability. They maintain that no test has enough reliability to warrant making decisions based solely or primarily on the basis of test scores.

Reliability is also compromised by the way the tests are administered. Teachers, in their efforts to help students score well, introduce variables into the system. Interventions include teaching to the test, teaching test-taking skills, allowing students extra time to complete the test, correcting answer sheets, and even asking low-performing students to stay home on test days (Gilman & Reynolds, 1991). Nolen, Haladyna, and Haas (1992) use the term “test pollution” to describe situations where test scores are distorted. They pose three classes of pollution: a) the ways that schools and teachers prepare for tests, b) variations in test administration conditions, and c) factors residing outside the influence of schooling that are known to influence test performance. In general, “test pollution” is the unequal and uncontrolled variety of test preparation and test administration that contaminate all comparisons among schools because students did not have the same opportunities to perform on the test. If the creators of the test normed the test with sample students whose test environment and administration was not polluted, then the scores of the students who did receive preparation will have different meanings (Urduan & Paris, 1994).

Validity Concerns

Validity of a test tells us whether a test measures what it claims to measure, how well it measures it, and what can be inferred from that measurement (Neill & Medina, 1989). They challenge the validity of standardized tests on several fronts.

For standardized tests to be valid, one must assume that the underlying trait being measured develops in a relatively consistent fashion among all individuals. They conclude this is not a valid assumption because of the individual differences among students. In addition, they assert tests rely upon outdated and inaccurate views of child development and therefore cannot accurately measure growth and change in students' knowledge, ability, or skills. In fact, Birkmire (1993) references Eckland's 1980 longitudinal study of the class of 1972. He concluded that performance on a test of basic skills did not correlate with post high school employment rates or income level. This calls to question the use of competency tests to predict future employee and societal benefits.

Neill and Medina (1989) reviewed research and concluded standardized tests compound biases against non-white students. They state tests tend to reflect the language, culture, or learning style of mid- to upper-class whites. Students of different ethnic groups, economic status, or English-proficiency therefore are disadvantaged when taking standardized tests. The tests also assume students share a common cultural experience, which is not a valid assumption. Finally, they point out that there are many ways of "knowing" and solving problems, but standardized tests do not take this into account.

These findings are confirmed by research of literature conducted by Peter Sacks (1997). He found that standardized tests generally have a questionable ability to predict one's academic success. He further found that standardized test scores tend to be highly

correlated with the students' socioeconomic class, substantiating Neill and Medina's findings. Sacks also recognizes standardized testing can reward superficial learning and can drive instruction in undesirable directions and thwart meaningful education reform. For these reasons, he states, "Teachers, researchers, and other educators have expressed widespread disenchantment with the results of several decades of standardized testing in the United States" (p. 27).

Teacher Attitudes toward Standardized Tests

Teacher attitudes have been studied in both high-stakes and low-stakes test environments. Gordon and Reese (1997) define high-stakes tests as "...standardized achievement tests used as direct measures of accountability for students, educators, schools, or school districts with significant sanctions or rewards attached to test results" (p. 345). As an example, Bushweller (1997) reports that principals and superintendents can be fired in Texas for poor school performance on standardized tests. In addition, school boards can be dissolved if districts score poorly. In Maryland, schools can forfeit thousands of dollars in reward money, or can even be taken over by the state, if scores are consistently low. In Florida, teacher evaluations are partially based on their students' standardized test scores.

Low-stakes environments are those where standardized tests are used, but they are not utilized as the primary measure of accountability nor have profound consequences for students and teachers (Soltz, 1992).

Soltz (1992) surveyed elementary teachers in two low-stakes school districts. His research question was, "Is there a relationship between teachers' perceptions of a standardized test, their administration of that test, and subsequent performance of their

classes on that test?" (p.4). In general, the answer was "no." Although a number of teachers expressed negative feelings toward standardized tests, the responses ranged from positive to negative. There was no consistent negative response. However, 62% of the teachers disagreed with the statement that the standardized tests seemed like a "fair measure" of what they taught and two-thirds of the teachers did not believe the mandated standardized tests supported district goals. Over one-half of the respondents reported they would like to learn more about the standardized tests, indicating they maintained an open mind about tests. The teachers reported they spent a lot of time preparing for the test, but the times involved were less than the times associated with test preparation in high-stakes environments. Overall, Soltz did not find any correlation between attitude, time spent in test preparation, or students' performance. Perceptions varied within and between grades, and were not as negative as reported in other research studies. He also found that teacher perceptions were generally unrelated to how they went about preparing for the mandated standardized tests.

Corbett and Wilson (1989) studied teacher attitudes in the low-stakes environment of Pennsylvania and the high-stakes environment of Maryland. Data collection included interviews with administrators, teachers, and students at twelve sites (six school districts in each of the states), a survey of central office administrators, principals and teachers from 207 Pennsylvania districts and 23 Maryland districts, and follow-up fieldwork.

In Pennsylvania, teachers reported that they had reservations on whether the tests improved the lives of either teachers or students. The stakes increased when the newspapers published the test scores. Teachers began to take the tests more seriously, but for political reasons, that is, the increased attention placed on the scores by parents,

administrators, and legislators. As the stakes were raised, the school districts' responses took on the flavor of a devotion to test scores.

In Maryland, teachers paid more attention to improving the test results. Teachers felt the tests generated a greater disruption of their work lives. They felt there was a decreased reliance on their professional judgment and they had a heightened concern about their liability.

For both states, the teachers found little justification for adding another test to the set of existing instruments being administered at the local level. Teachers perceived that statewide testing programs offered relatively low benefits for students because the tests did not provide information the schools already possessed through other means.

There have been several studies conducted in high-stakes states to assess teacher attitude. The studies looked at teachers who taught at different levels of instruction. Overall, the responses to attitude surveys showed teachers had a negative perception of standardized tests.

Green and Stager (1986) surveyed teachers in Wyoming to measure their attitudes toward standardized tests. They posed four hypotheses: a) there are significant differences in attitudes toward testing among teachers who have zero, one, and two or more courses in tests and measurement; b) there is a significant positive correlation between favorability of personal past experience with tests and attitude toward testing; c) men have more favorable attitudes toward testing than do women; and d) there are significant differences in attitude toward testing across grade levels, with attitudes being more positive at higher grade levels. The results showed the first hypothesis was not supported, but the remaining hypotheses were supported.

In general, teacher attitudes toward testing ranged from indifference to negative feelings. Teachers rejected the idea that their salaries should be based on standardized test scores. They thought scores were inappropriate measures for evaluating teacher effectiveness. The teachers generally were favorable toward using competency tests for students. They viewed standardized tests as serving some purpose, but did not see them as useful as classroom tests.

Nolen et al. (1992) surveyed teachers and administrators in Arizona to determine: a) how teachers and administrators used standardized test scores, b) how they perceived peers and others used test scores, c) how schools prepare students to take the tests, d) conditions under which the tests were administered, and e) teacher and administrator views of the effects of the testing program on students, teachers, and administrators. Nearly 2,000 teachers and 600 administrators responded to the survey. The surveys contained a “fairly exhaustive” list of possible uses for test scores. Respondents indicated the scores were not routinely used for most of the purposes provided on the list. The exception was to use the scores to identify students for gifted or remedial programs. And, about 45% of the teachers and administrators use standardized test scores to evaluate class or school effectiveness. Both teachers and administrators overestimated the extent the other group used test scores for the listed purposes. Teachers reported they felt the tests had many effects on students, including a variety of symptoms that are likely to be related to anxiety. In general, there was evidence that discomfort with the testing program may be due to the belief that the testing process has a negative effect on students and teachers. Over two-thirds of the elementary personnel reported that teachers feel threatened by the results of the tests. They were frustrated because of their doubts about

the usefulness of the tests themselves. Over 80% of the respondents felt that the state-mandated annual achievement testing was not worth the time and money spent.

Urdan and Paris (1994) surveyed K-8 teachers in Michigan. The purposes were to (a) examine teachers' attitudes, perceptions, and practices regarding standardized achievement tests and (b) examine ways in which teachers differ in their attitudes, perceptions, and practices regarding standardized tests (that is, to see if they differ according to the number of years of teaching experience, grade level taught, and achievement level, race, and socioeconomic status of children taught).

Similar to the Arizona study, 77% felt the tests were bad and not worth the time and money spent. Less than 10% felt the tests reflected what the students learn in school. Only 5% felt the tests were accurate for minority students and students with limited English language skills. Only 21% felt their school districts provided adequate training for the administration of standardized tests and interpretation of the resulting scores. In general, the respondents had negative views toward standardized tests and did not believe the tests were good for education. The researchers did not find appreciable differences in the responses of the various teacher groups.

Gordon and Reese (1997) surveyed over 100 teachers in Texas and interviewed some of them to gather in-depth qualitative data on perceptions of teachers in public schools regarding the effect of the Texas Assessment of Academic Skills (TAAS) testing program on students, teachers, and schools. Concentrating on the reported effects on teachers, they reported that as more emphasis is placed on TAAS-related course content, the less they emphasize other content in their classes. Teacher stress increases as the TAAS testing time approaches. They express concern, frustration, and disappointment

when at-risk students perform poorly on the TAAS. They feel they are held accountable for TAAS-related content, but are not being held accountable for effective teaching.

In general, they thought TAAS does not encourage or measure higher-order thinking skills. The TAAS program makes them less innovative and lessons are less interesting. The interviewees voiced the belief that TAAS scores are not accurate measures of the academic progress their at-risk students have made.

Birkmire (1993) surveyed teacher attitudes toward the Ohio high school proficiency test. In general, the surveyed teachers agree that students should have a minimal competency, but questioned whether the state competency test should be used as the only determinant for whether a student should receive a high school diploma. The teachers indicated the competency test increased both demands on their time and the amount of paperwork required for their jobs. The teachers were negative toward state legislators questioning the abilities of teachers based on competency test scores. Finally, the teachers did not see the attitudes or grades of their students as improving as a result of the state competency testing program.

Arnott (1992) analyzed Ohio elementary teacher attitudes toward standardized achievement tests for her Master's Thesis at the University of Dayton. She surveyed teachers in two Ohio school districts. The majority of respondents had reservations regarding the reliability of standardized tests and questioned the validity of standardized tests. Most respondents felt standardized tests were not uniformly administered. Respondents were divided over how much standardized test results should be used in developing and refining their course work. They also felt they were not adequately trained to administer the tests. Over 90% of the respondents did not feel parents understood

standardized test scores and 60% thought administrators did not use test scores in a wise manner.

Smith (1991b) interviewed a number of teachers in Arizona for a qualitative study about attitudes toward standardized tests. She discovered teachers had negative emotions as a result of the publication of test scores. The publication of test scores produced feelings of shame, embarrassment, guilt, and anger. Teachers indicated they do what's necessary to avoid low scores. The teachers felt the tests were invalid, but still felt the necessity to raise scores. This brought on feelings of dissonance and alienation. Finally, they felt the emotional impact of testing on small children generated feelings of guilt and anxiety among teachers.

Salmon-Cox (1981) interviewed 68 elementary teachers on a variety of subjects, but included questions about teaching style and evaluation. She found that most teachers rely on a number of assessment approaches, and only 3% mentioned standardized tests as an assessment technique. The teachers said standardized test information was used as a supplement to or confirmation of information they already had about individual students. Therefore, they felt the tests lacked importance. The researcher concluded that teachers want diagnostic tests that are precise, are closely matched to curricula and instruction, and are timely. Achievement tests of the kind widely used do not match these criteria.

Research by Sacks (1997) looked at a number of studies and papers and provides a summary of teacher attitudes toward standardized tests. Standardized tests generally have a questionable ability to predict academic success, especially for certain subgroups of students. Standardized tests tend to be highly correlated with students' socioeconomic status. Standardized tests can reward superficial learning, drive instruction in undesirable

directions, and thwart meaningful educational reform. Teachers for K-12, especially, testify that standardized tests do not accurately measure students' abilities and that practices of "teaching to the test" render the test scores virtually meaningless.

Impacts of Standardized Testing on Instruction

The literature revealed that many teachers spend class time preparing students for standardized tests. Nolen et al. (1992) noted the time teachers spent in preparation ranged from no time spent to spending two or more months of daily or weekly test preparation activities. In addition, 25% of the elementary teachers responding stated administrators required them to spend time preparing their students.

Corbett and Wilson (1989) found teachers felt pressure to improve scores and felt forced to take time to teach to the test. One teacher reported she spent from September to November on basic (reading) skills rather than on developmental programs. And, one teacher is quoted as reporting "It takes too much time...too much of that time has to be taken from other stuff I used to do." Likewise, in interviews with teachers, Smith (1991b) was told by one teacher that part of every day was spent on the district study skills manual. Another complained that test preparation took a 100 hour bite out of her instruction time. Smith (1991a) also had one teacher report that she spent 80% of her teachable time in the three weeks prior to testing on administering work sheets from district materials.

In their study of high stakes testing in Texas, Gordon and Reese (1997) found even more time impact. Some teachers planned and taught TAAS preparation the entire year. Some teachers held special TAAS practice sessions daily; others gave TAAS practice quizzes or tests weekly, every six weeks, or every semester. One teacher reported

she spent nine days each semester on TAAS preparation. Many felt that the intensity of test preparation increased in the four to eight weeks prior to the TAAS.

Urduan and Paris (1994) asked teachers how many hours they personally spend preparing students to take standardized tests, when they start preparing students for the test, and in the months before the test, how often their classes prepare for it? The responses were aggregated to form a single reliable scale. Forty-two percent responded to the high end of the scale.

Gordon and Reese (1997) found 92% of the teachers they interviewed reported they often or always give students help as they prepare to take standardized tests. The time spent preparing students for standardized tests involves numerous activities. These activities have a very real impact on the instructional decisions teachers make. Many of the same effects on instruction were echoed among the studies. Following is a list of effects on instruction identified in the studies:

Teaching general test taking skills or how to properly mark test answer sheets (Bushweller, 1997; Gordon and Reese, 1997; Nolen et al., 1992; Smith, 1991a, 1991b; and Urduan and Paris, 1994)

Teaching or reviewing topics covered by the upcoming test (Gordon and Reese, 1997; Madaus, 1991; Nolen et al., 1992; Smith, 1991a)

Teaching basics instead of higher order thinking skills or problem solving (Birkmire, 1993; Madaus, 1991; Neill and Medina, 1989; Sacks, 1997; Stake, 1991)

Informing parents and students about the importance of proper rest, nutrition, and attendance (Nolen et al., 1992, Smith, 1991a, 1991b)

Encouraging students to try hard (Smith, 1991a, 1991b; Urduan and Paris, 1994)

Spending a lot of time on drill and practice of test topics (Gordon and Reese, 1997; Madaus, 1991; Smith, 1991a)

Using commercial test-preparation materials (Gordon and Reese, 1997; Nolen et al., 1992; Urdan and Paris, 1994)

Using the standardized test format in classroom tests and activities (Gordon and Reese, 1997; Nolen et al., 1992)

Teaching actual items from current tests (Nolen et al., 1992; Smith, 1991a)

Giving special remediation to students who had not done well on previous standardized tests or on practice tests (Gordon and Reese, 1997)

Gordon and Reese (1997) and Smith (1991b) reported some teachers go so far as to delete entire subjects from their curricula so they can concentrate on preparing students for standardized tests. Teachers in Texas reported to Gordon and Reese that classes in non-TAAS content areas can virtually disappear in the weeks ahead of the test administration. Likewise, in one Arizona classroom, Smith found that science went from laboratory/hands-on instruction in the fall to no science at all in the spring as test time approached; in another class, social studies disappeared.

As stated before, some feel these test preparation strategies invalidate the usefulness of standardized test scores. But, as Brandt (1989) quotes Madaus, "When the stakes are high, people are going to find ways to have test scores go up" (p.25).

Summary

Standardized tests have been used in the United States for over 100 years. The use of these tests is increasing as a part of educational reform movement to hold teachers and schools accountable for what students learn. Some educators question the validity and

reliability of standardized tests because of the variations in test preparation that teachers provide. Studies show teachers generally have negative feelings toward standardized tests and spend a lot of instructional time preparing their students for the tests. The next chapter describes the subjects and setting of the study and explains the data collection and treatment.

CHAPTER III

METHODS

The purpose of this study was to determine the relationship between the attitudes of teachers, in a selected elementary school, toward state-mandated standardized tests and how the standardized tests affect their instructional decisions. I quantified the attitudes (within a range of positive and negative attitudes) and the degree of effect standardized tests have on instructional decisions. My hypothesis was that there is a positive correlation between the degree of negative attitude and the amount of effect on instructional decisions.

Subjects

The subjects were elementary teachers employed at Hutchison Elementary School, located in Herndon, Virginia. Eighteen of the subjects were female and two were male. Their ages ranged from 24 to 59 years old. The subjects' years of teaching experience ranged from a first year teaching to 29 years of experience. Together, the subject teachers had experience teaching grades kindergarten through sixth grade, as well as physical education, art, reading, emotionally disturbed children, and learning disabled children.

Setting

Hutchison Elementary School is in a suburban setting. It is part of the Fairfax County Public School System. Fairfax County is located in northern Virginia, within the Washington, D.C. metropolitan area. The local neighborhoods represent the working class and middle class. There is an extensive variety of ethnic backgrounds among the students, with over 30 countries represented in the student body.

Data Collection

A combination of surveys and interviews was used to ascertain teacher attitudes toward standardized tests and how much affect standardized tests have on their instructional decisions. The surveys were based primarily on questions developed by Arnott (1992) for her Master's project for the University of Dayton and used a Likert scale for respondents to indicate their answers. The survey is at Appendix A. The survey was field-tested with five educators, including a former elementary school principal, three elementary school teachers, and a ninth grade teacher. One question was added to the survey per a suggestion from the ninth grade teacher and several questions were modified to improve their clarity. The purpose of the interview was to gain additional insight into the teacher attitudes towards standardized tests and if or how the use of standardized test affected their instructional decisions. The guide used for the interviews is at Appendix B.

The state of Virginia mandates that elementary students in the second, fourth, and sixth grades take the Stanford 9 test in the fall. Teachers of those grades were surveyed to determine their attitudes towards standardized tests. In the 1998-1999 school year, the state implemented new tests called the Standards of Learning (SOL), to be taken by all students in the third, fifth, and eighth grades and in high school. The SOL will be used as a tool to hold schools accountable for students' performance, as well as for student graduation requirements and school accreditation. Therefore, the SOL is a high-stakes test, which may directly impact instructional decisions. The fifth grade SOL covers Virginia history, which is taught in the fourth grade, and schools have the option of administering that portion of the SOL to fourth grade students. Since the subject school does administer the Virginia history SOL to the fourth grade students, teachers of grades

three, four, and five were given the survey to determine their attitudes and were also interviewed to assess the impact the SOL has on their instructional decisions.

Treatment

The Fairfax County Public Schools Department of Test and Evaluation was contacted to obtain permission to conduct research at Hutchison Elementary School. Since the research was restricted to one school, only the permission of the principal of the school was required. The principal was provided with a written request, outlining the research and including the proposed survey and interview guide. He supported the research and invited the investigator to present the project at a staff meeting. The principal provided a list of the school staff to aid in addressing correspondence.

Concurrently, information was sent the University of Dayton Committee for the Protection of Human Subjects for review. The package included the research proposal, the proposed survey, the proposed interview guide, and the proposed informed consent form. The Committee reviewed the proposal and gave approval to proceed.

Using the in-school teacher mailboxes, a letter was sent to 40 teachers asking them to participate in the research. The letter included two copies of the informed consent form, with instructions to sign both forms and to return one to the investigator. The teachers included all second through sixth grade classroom teachers, as well as the physical education, music, art, English as second language (ESL) teachers, and teachers for the learning disabled (LD). Twenty-five teachers (65% return rate) returned signed consent forms. Each consenting teacher was provided with a survey, again using the in-school mailboxes. Reminders were sent to those teachers who did not return the survey by the deadline. Unfortunately, the survey did not include a code or other means for the

teachers to identify themselves, so another letter was sent requesting they include a code number. Twenty-two surveys were returned (88% return rate), of which two teachers could not be identified. These two surveys were not included in the data. Table 1 shows the subject response by grade level.

Table 1

Subject Response by Grade Level

Grade	Consent	Survey	Interview
2nd	5	3	-
3rd	6	5	5
4th	5	5	5
5th	3	2	2
6th	3	2	-
LD	2	2	-
Art	1	1	-

Note. A dash indicates interviews were not required.

Once the surveys were tallied, the investigator personally contacted all third, fourth, and fifth grade teachers who returned surveys and scheduled interviews.

Interviews lasting from 15 to 30 minutes were conducted with each subject teacher, with hand-written notes taken to record their response.

Summary

This chapter explained the methodology used to complete the study. It described the subjects, setting, and the data collection procedures. It also discussed the questionnaire survey and interview instruments and expressed the return rate. Chapter IV describes the results of the study, including both survey and interview

techniques, with some using the school counselor as a resource. Table 3 shows their assessment of the time spent on test preparation activities.

Table 3

Time Spent on Test Preparation/Review

Grade	Subject	Time Spent
3rd	All	50% each day for 4 weeks prior to test
4th	Teacher 13	20 - 30% each day; 65% of time for 3 weeks prior to test
	Teacher 14	30 - 50% each day for 4 weeks prior to test
	Teacher 15	80% of social studies time; 30% of day spent integrating social studies and language arts
	Teacher 16	15% each day for 4 weeks prior to test
	Teacher 17	50% each day for 4 weeks prior to test
5th	Teacher 18	50% each day for 8 weeks prior to test
	Teacher 21	20% of time for 12 weeks

The third grade teachers set up a class rotation for SOL review. For four weeks, one-half of every day was spent in review. Each teacher was responsible for one subject area. The classes would rotate among the teachers each day for the review periods. In addition, the individual teachers incorporated review in their own classroom instruction. Four of the third grade teachers indicated they started test preparation at the beginning of the year, planning the curriculum around the test to ensure all testable material was covered. The fifth teacher indicated she spent 30 - 40% of class time on test preparation, with preparation incorporated into the classroom instruction.

All five of the fourth grade teachers indicated they begin test preparation activities at the beginning of the year. Both of the fifth grade teachers said they incorporate SOL test preparation activities into their curriculum. One fifth grade teacher stated "I don't teach anything not SOL related." She also identified students that needed additional help to succeed on the test and sent them to SOL coaches and resource teachers. This group spent anywhere from 15 to 40 minutes per day for four to six weeks learning test-taking strategies and other skills.

The implementation of the SOLs has impacted all of the teachers' instructional decisions. They all have eliminated or cut back on some classroom activities or subject areas so that they have time to cover the testable material. (One teacher said she saves the "fun" things for after the test.)

In addition, three of the fourth-grade teachers said they don't have time to cover some things to the depth they would like. Time constraints force them to cover the testable material at the expense of depth of understanding. Four grading periods worth of material must be covered in three grading periods. For example, the fourth grades have to complete the entire Virginia history textbook and the fifth grades have to complete the entire math textbook by mid-May when the test is given. One fourth-grade teacher and one fifth-grade teacher said they don't always have time to make sure all of the students fully understand before moving on. The teachers felt it was important for all of the students to at least be exposed to all of the testable material, even if some students do not fully understand the material.

Finally, the SOLs have changed the way the teachers teach. Virtually all reported they integrate the testable material into their curriculum, with most specifically stating

they begin test preparation at the beginning of the year. All of them commented on the quantity of testable material to be covered and the limited amount of time available. And, more time is being spent on drills and memorization at the expense of hands-on, creative activities.

Summary

This chapter described the results of the responses to the surveys and interviews. Responses to each survey question were shown. Information from the interviews pertaining to time spent on test preparation and other impacts on instructional decisions were presented. Chapter V provides an analysis of the data and shows the relationship between the subject teachers' attitudes towards standardized tests and the impact that standardized tests have upon their instructional decisions.

CHAPTER V

DISCUSSION

Survey Results

The majority of the teachers surveyed indicated they did not feel standardized tests were reliable or valid. Some teachers indicated they have concerns with standardized tests because of the large number of immigrant, ESL, transient, and impoverished students at Hutchison. Nolen et al. (1992) recognized that test reliability and validity are dependent upon a uniform student base, which Hutchison does not have. Therefore, the teachers' concerns appear to be well-founded.

An overwhelming majority of the teachers surveyed indicated that classroom achievement is more important than standardized test scores. Neill & Medina (1989) recognized individual differences among students affect how they perform on standardized tests. At Hutchison, many of the immigrant and ESL students do not have a strong foundation in United States schooling. Therefore, they are at a disadvantage when taking the standardized tests. The teachers, however, are in a position to evaluate the students on an ongoing basis throughout the year. This ongoing evaluation appears to be of more use to the teachers than standardized test scores. A slight majority (53%) did believe that student standardized test scores are useful. However, a majority indicated teachers are not provided test scores in a timely manner. The SOLs are administered in May. The scores are not published until after school dismisses for the year. As Jaeger (1991) points out, one valid use of test scores is to assist teachers in identifying areas in which students need help. Since scores are available only after the students leave the classroom, this is not possible. Over one-third of the teachers surveyed indicated that

teachers did not make use of test scores in a wise manner, but over one-third were undecided. This is consistent with Resnick's (1981) findings. He found standardized tests were not relevant to teachers because the tests only measure the teachers' cognitive goals for students and none of the social goals.

A majority of the teachers surveyed felt they were given adequate instruction on how to administer the standardized tests. They were provided with guides and training to ensure uniform administration. Because of the guidance, a majority felt standardized tests are uniformly administered in the school system. This addresses one of Nolen et al. (1992) possibilities of test pollution, notably variances in test administration.

A majority of the teachers surveyed indicated teachers teach materials they anticipate will be on the SOLs. This was confirmed during the interviews. All of the teachers interviewed said they spend a lot of time covering testable material, often at the expense of other subjects or activities.

Nearly half of the teachers surveyed were undecided if test score abuse occurs. This may be because the survey did not include a precise definition of test score abuse. Thirty-seven percent agreed that score abuse occurs. Sproull & Zubrow (1981) recognized four uses of scores: diagnosis and placement of individual students, evaluating instructional programs, as a measure of student end-of-year achievement, and reporting to outside audiences (parents, legislatures, etc.). Jaeger (1991) also recognized standardized tests could be used as an assessment of basic skills. Virginia is using test scores as accountability tools to measure school performance. The scores are used to determine school accreditation, as a basis for providing schools and staff with monetary rewards and

sanctions, and for student credit towards graduation. None of these conform to Sproull and Zubrow's or Jaeger's concepts of recognized uses of scores.

A majority of the teachers surveyed indicated that schools should not be compared with one another on the basis of their standardized test scores. In Virginia, school scores are published in newspapers, where parents and administrators can compare schools. In Fairfax County, the School Board looked at the schools' SOL scores and identified 20 elementary schools with low scores. Those schools shared common traits: student bodies with high numbers of immigrant, ESL, transient, and poverty-level students. The School Board implemented Project Excel for these 20 schools. Project Excel allocates additional resources to the schools to assist in raising the SOL scores. In addition, the Project Excel schools are being evaluated on a yearly basis to determine if they are improving. Schools that exceed performance goals receive cash bonuses for teachers and staff; schools that do not improve receive additional help. While the bonuses or sanctions are based on comparing each individual school's annual scores with its previous scores, it is likely that the schools will be compared to each other and the rest of the schools in the county.

One-hundred percent of the teachers surveyed indicated that teachers are not involved in selecting standardized tests used in the school. In Virginia, the State mandates the tests used. In 1994, the Virginia Department of Education contracted with four school districts to develop standards of learning in mathematics, English, science, and history/social studies. Over 5,000 Virginians, including parents, teachers, professional organizations, the business community, and interest groups were involved in the process of evaluation and development of the standards. Tests of the standards were developed in a partnership with Virginia educators (including teachers), Harcourt Brace Educational

Measurement, and the Virginia Department of Education. Test items were field tested in 1997 and the first round of testing was administered in 1998 ("A Better Education for a Better Future," 1999). While some teachers were involved with designing the test, the State chose the tests and mandated their use.

There was no clear consensus among the teachers surveyed as to whether teacher classroom performances are evaluated on the basis of their students' test scores. Officially, the test scores are being used to evaluate schools (not teachers) and hold them accountable for the students' performances. However, within the Project Excel schools, principals are being given more authority on hiring and transferring staff. Also, teachers in Project Excel schools may earn monetary rewards if the school's test scores exceed its goals. While the test scores are not yet officially being used to evaluate teacher performance, it is understandable that some teachers feel others are evaluating their performance based on their students' test scores.

A majority of the teachers surveyed indicated that they felt parents do not understand standardized test scores. Hutchison has a large immigrant and transient student population, with many parents who speak little or no English. Lack of English skills may be one barrier to parental understanding of the test scores.

A majority of the teachers felt they receive adequate training in the subject areas. Course curriculums have been aligned with the test content. The State and County provide guides and resources related to SOL content. However, some teachers interviewed said they do additional research into content area and use other resources to ensure all material is covered. Because of the time spent on SOL content area, there is strong evidence that the majority of teachers "teach to the test."

There was no clear consensus among the teachers surveyed as to administrators making wise use of standardized test scores. Forty-two percent indicated administrators did not make wise use, but thirty-seven percent were undecided. Implementation of the SOLs is a new phenomenon, and the use of the test scores is being reviewed by state and county administrators. Since the system is still in a state of flux, it is understandable that the teachers are unclear about how the test scores are being used.

A vast majority of teachers surveyed indicated they felt there was too much emphasis on standardized tests. The SOLs are high-stakes tests, with school accountability, resource allocation, accreditation, student graduation, and school rewards and sanctions at issue. Because of the high-stakes nature of the tests, all of the teachers have had to adjust their teaching to foster student success on the tests. Often, this has been at the expense of non-testable subjects such as health.

The teachers' answers to the survey can be interpreted as representing a positive or negative attitude towards standardized tests and their use. For instance, an answer of "disagree" or "strongly disagree" with the statement "In general, standardized tests are reliable (the results of what the tests measure are consistent across the student population)" indicates a negative attitude towards standardized tests. Table 4 shows how answers for each question are categorized as either negative attitude or positive attitude. An "undecided" answer is treated as neutral.

Table 4

Assessment of Attitude

For Questions	1,2,4,8,10,12,13,15,16	an answer of "disagree " or "strongly disagree"	indicates a negative attitude
		an answer of "agree" or "strongly agree"	indicates a positive attitude
For Questions	3,5,6,7, 9,11,14,17	an answer of "agree" or "strongly agree"	indicates a negative attitude
		an answer of "disagree" or "strongly disagree"	indicates a positive attitude

Based on this interpretation of answers, the results of the survey show a marked negative attitude towards standardized tests. Only three teachers showed a more positive than negative attitude. Table 5 shows the number of responses in each category by teacher.

Table 5

Analysis of Attitude by Teacher

Teacher	Negative Responses	Positive Responses	Undecided Responses
1	11	5	1
5	7	4	6
7	5	11	1
8	10	2	5
10	9	2	6
11	14	0	3
12	9	6	2
13	8	5	4
14	7	5	5
15	11	6	0
16	6	7	4
17	13	2	2
18	12	4	1
21	10	5	2
23	12	4	1
26	11	6	0
27	6	6	5
28	5	8	4
29	11	5	1
39	7	6	4
Total	184	99	57

Attitude does not seem to be a factor of teaching experience. Figure 1 shows the number of positive, negative, and undecided responses plotted against individual teachers' years of teaching experience. The three teachers with 25 years of experience had more positive than negative answers; however, teachers with more than 25 years of experience had more negative answers. Therefore, negative attitude cannot be attributed to the amount of teaching experience.

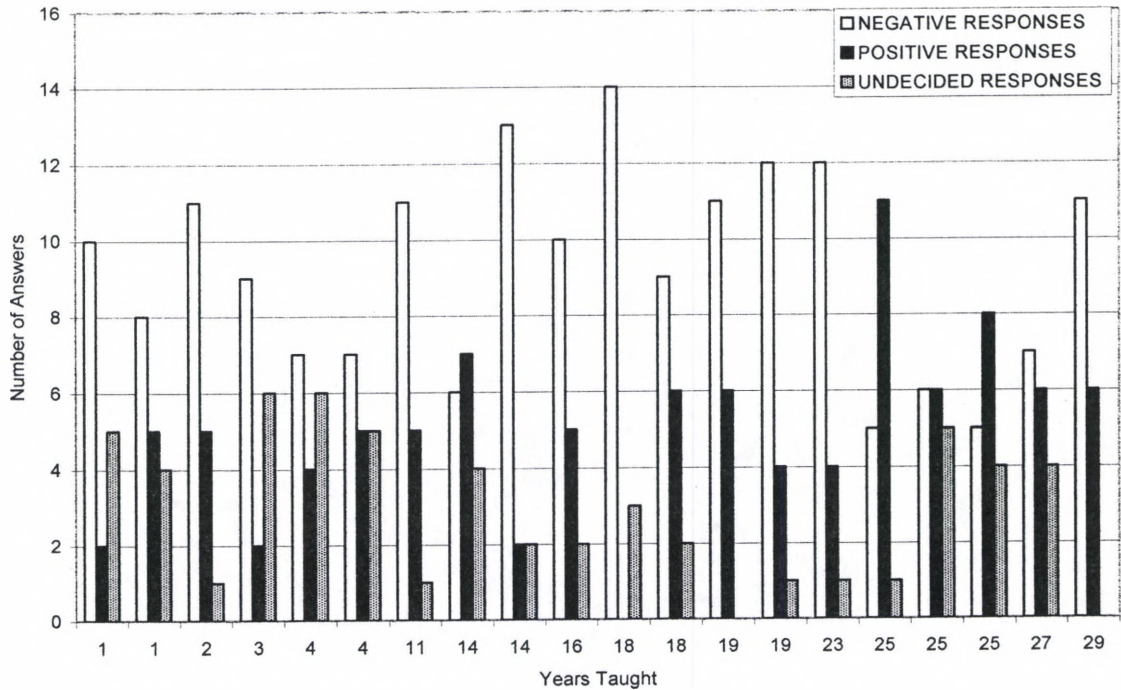


Figure 1. Attitude by Years Taught

There is no relationship between grade level and attitude. Table 6 shows that all of the grades had more aggregate negative responses than positive responses. Regardless of grade or subject taught, the overall attitude of the subject teachers was negative.

Table 6

Teacher Attitude by Grade

Attitude	Grade 2	Grade 3	Grade 4	Grade 5	Grade 6	Other
Negative	23	43	45	22	19	28
Positive	17	21	25	9	10	17
Neutral	11	17	15	3	5	6

Impact on Teaching and Instructional Decisions

The implementation of standardized tests, specifically the SOLs, has had a major impact on the teaching and instructional decisions of the teachers at Hutchison Elementary School. All of the teachers interviewed have made adjustments in their classrooms to accommodate the tests. Table 7 shows activities and/or subject areas that teachers in grades three through five have reduced or eliminated due to the implementation of the SOLs. According to the teachers surveyed, the most serious impact is the time-pressure the teachers feel. They indicated that there is so much material to be covered that there is not enough time to cover the subject matter in depth. Since they have to spend so much time covering the testable material, they feel there is not enough time for fun, creative activities or hands-on activities.

Table 7

Activities Eliminated or Reduced

Teacher	Activities
7	Cut poetry units; hands-on science; some research papers; fable and fairy tale units; crafts and cooking as part of culture; some health; gifted and talented projects
8	Cut back on reading time; health and family life education; cut back on science and hands-on science projects
10	Cut "fun units" like unit on butterflies
11	Cut fun activities like fairy tales; cut time spent on presidents and health
12	Cut back health; fewer creative crafts and art activities
13	Cut time on science
14	Cut five themes of geography and how geography affects people and government; cut other themes of government and economics
15	Shortened writing program; modified math program, spending less time on solving strategies; cut in-depth study of colonial Virginia; cut some crafts and art projects tied to Social Studies
16	Hardly any health covered; "shirked" writing program
17	Cut back emphasis on writing; not as much time spent on editing and rewriting
18	Cut back a lot of creative things; critical thinking activities; classroom newspaper; middle ages and Renaissance in Social Studies; art activities
21	Social Studies not covered in detail; cut back on creative projects like research

Implementation of the SOLs has had an impact on other instructional decisions of teachers in grades three through five. Table 8 shows the individual teacher's assessment of the impact of SOLs on their instructional decisions. It can be argued that the impact is neither positive nor negative. School systems routinely update and change the curriculum and programs of study, requiring teachers to update their instruction. For the SOLs, the teachers appear to have seamlessly incorporated SOL instruction and test preparation into the classroom routine. In the interviews, several teachers indicated that they incorporate SOL concepts into other subject areas like reading, writing, math, and science. This may provide the students with a more cohesive approach to learning the material. All of the teachers interviewed said they use drills and games for subject matter review. Since drills and reviews are a normal part of classroom instruction, the SOLs may have merely shifted the emphasis of the drills to a more focused area.

There are some negative aspects, however. Several teachers indicated they use more drill and memorization activities and fewer creative activities. Also, four teachers said they do not have time to make sure all students understand material before moving on. They expressed real concern that some students are being "left behind" because there is so much material to be covered for the test.

Table 8

Impact on Instructional Decisions

Teacher	Impact
7	Executed planned units in classroom; test preparation throughout the year
8	Spent most of time covering testable material
10	Slowed down; did not spend as much time covering things in all subject areas; spent less time making sure students understood
11	Incorporated SOL concepts into reading, writing, and science
12	Put year's worth of curriculum into 27 weeks instead of 36
13	SOL has had impact on how conducts class - not positive or negative
14	"SOLs drive my instruction to a certain degree and I don't like it." Would like to spend more time on process than on covering specific facts
15	Great deal of preparation and in-class time-on-task spent on SOL preparation. Integrated language arts with social studies. SOL has had no effect on style, but requires more teacher research and materials
16	Used to take children as far as they could go. Now must teach what the State and County says. Not time to handle both and cover the curriculum effectively
17	Created games, activities, and review sheets based on resource guide provided by the state and additional research
18	Don't teach anything not SOL- related. Wants to know if kids are ready to move on, but there's too much material to cover
21	Direct impact from SOL: lots of drill and memorization; less creative thinking.

Relationship Between Attitude and Impact on Instructional Decisions

A direct comparison between attitude and impact on instructional decisions is difficult to make because the interview responses cannot be quantified like the attitude assessment survey. However, a general relationship can be shown between attitude and the time spent on test preparation. Since the teachers conducted test preparation activities for different lengths of time, it is difficult to draw direct comparisons between them. In order to establish a baseline, it was assumed that the teachers began preparation eight weeks prior to the test. All percentages of time spent were converted to the eight-week baseline. For instance, if a teacher indicated she spent 50% of her class time per day for four weeks in test preparation, the eight-week value would be 25% of class time. Table 9 shows the time (raw and adjusted percentages) the teachers spend on test preparation compared to their attitude scores.

Table 9

Comparison Between Time Spent on Test Preparation and Attitude Scores

Teacher	Time Spent on Test Preparation	Adjusted Daily Percentages	Negative Responses	Positive Responses
7	50%/day for 4 weeks prior to test	25% for 8 weeks	5	11
16	15%/day for 4 weeks prior to test	7.5% for 8 weeks	6	7
14	30 - 50%/day for 4 weeks prior to test	25% for 8 weeks	7	5
13	20 - 30%/day; 65% of time for 3 weeks prior to test	33% for 8 weeks	8	5
10	50% of time for 4 weeks; 30%/day integrating SOL material into curriculum	25% for 8 weeks	9	2
12	50% of time for 4 weeks; plus additional reviews	25% for 8 weeks	9	6
8	50% of time for 4 weeks	25% for 8 weeks	10	2
21	20% of time for 12 weeks	30% for 8 weeks	10	5
15	80% of social studies time; 30% per day spent integrating social studies and language arts	30% for 8 weeks	11	6
18	50% per day for 8 weeks prior to test	50% for 8 weeks	12	4
17	50% per day for 4 weeks prior to test	25% for 8 weeks	13	2
11	Integrated throughout year	N/A	14	0

Figure 2 shows the percentage of time each teacher spent on test preparation compared to the number negative responses he/she gave. Likewise, Figure 3 shows the percentage of time each teacher spent on test preparation compared to the number of positive responses he/she gave.

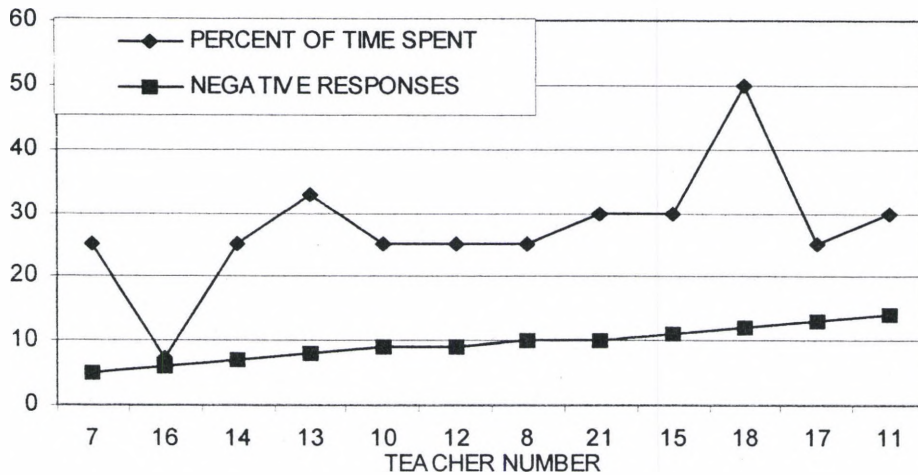


Figure 2. Percent of Time Spent Compared to Negative Responses

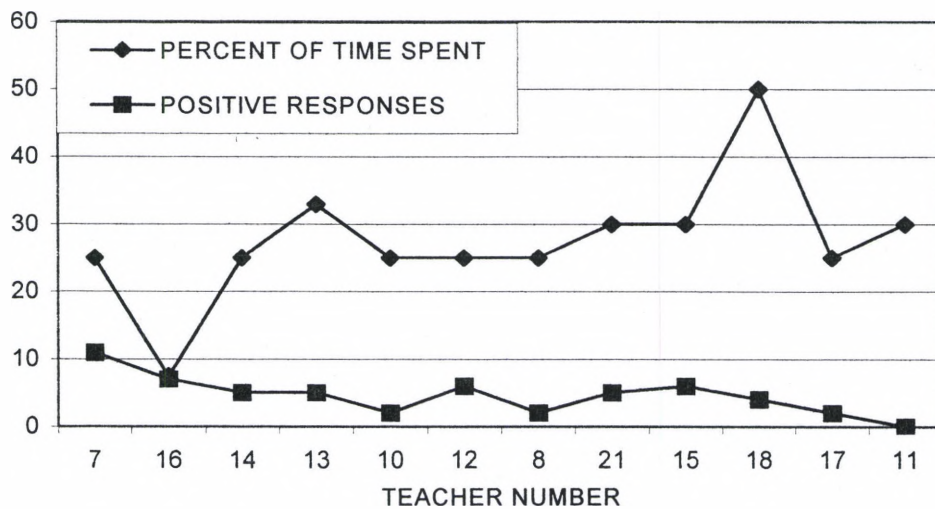


Figure 3. Percent of Time Spent Compared to Positive Responses

Figure 2 shows that as the number of negative responses increases, there is a slight increase in the adjusted percent of time spent in test preparation. Figure 3 shows that as the number of positive responses decreases, there is a slight increase in the adjusted percent of time spent in test preparation. Therefore, there is a slight but measurable positive correlation between negative attitude and the amount of time the subject teachers spent in preparing students to take the SOLs, and a slight but measurable negative correlation between positive attitude and the amount of time spent in test preparation.

Time spent on test preparation is only one measure of the impact that standardized tests have had on instructional decisions. Eliminating or reducing time spent on other activities and subjects and incorporating SOL instruction into the classroom curriculum were mentioned by virtually all of the teachers interviewed. However, it was not possible to draw a comparison between these areas and attitudes due to the variety of comments expressed by the teachers.

Summary

The purpose of this study was to analyze elementary teachers' attitudes towards the use of standardized tests and to determine the relationship, if any, between the attitudes expressed and the impact on instructional decisions.

A Likert-style survey consisting of seventeen questions was given to twenty teachers at a suburban Virginia elementary school to assess their attitudes towards standardized tests. Twelve teachers of grades three through five were interviewed to determine how the use of standardized tests, specifically the Virginia SOLs, affected their instructional decisions.

The survey responses showed the majority of teachers have a negative attitude towards standardized tests. Their responses and concerns are consistent with studies and research conducted in other high-stakes testing environments.

All of the teachers who were interviewed said the SOLs have had a major impact in the classroom. All of them spend considerable time in special test preparation activities, including "teaching to the test," test taking skills, games and drills, and relaxation techniques. All of the teachers interviewed have modified their teaching, incorporating SOL material into other subjects, with some spending more time on memorization and drills. In addition, SOL preparation requires so much time that the teachers must reduce time spent on creative activities or on some subject areas, and some teachers have even eliminated subjects such as health.

Comparing the attitude surveys with the interview results did show a relationship between the attitudes expressed and the amount of time spent preparing students to take standardized tests. In general, the more negatively a teacher felt about standardized tests, the more time was spent on test preparation.

The research findings are consistent with the previous studies and analyses reported in the review of literature and support the hypothesis.

Conclusions

The implementation of standardized tests, specifically the Virginia SOLs, has had a major impact on elementary classroom teachers. At Hutchison Elementary School, teachers in grades three through five indicated they have greatly modified their teaching and instructional decisions due to the SOLs. Entire subjects have been dropped in some classrooms. Other subjects and activities have been cut back so that additional time may be spent preparing students to succeed on the SOLs.

The SOLs were developed as a means to hold schools accountable for their students' learning. While the SOLs do provide measures of core knowledge, measures of other learning skills such as creative thinking and problem solving are not addressed.

The teachers feel pressured by the time necessary to prepare students for the test. This has greatly impacted their classroom instruction, and their attitude. One fourth grade teacher summed up her feelings when she stated "SOLs drive my instruction to a certain degree and I don't like it."

The perceived benefits of standardized tests need to be weighed against the loss of autonomy in the classroom and the increased stress felt by both teachers and students. Further analysis is warranted to determine if the trade-off is worth it.

APPENDIX A

Standardized Test Attitude Questionnaire

Several standardized tests are used in Fairfax County Public Schools. These include Stanford 9, Standards of Learning (SOL), and others. The following questions seek to determine your personal opinion of standardized tests in general and NOT just SOLs.

Demographic Information

Sex: M or F (circle one)

Age: _____

Grade(s) taught _____

Years of teaching experience: _____

Directions: Read each statement. Circle the response that most accurately reflects your attitude toward the statement. The responses are: SA=Strongly Agree, A=Agree, U=Undecided, D=Disagree, and SD=Strongly Disagree.

1. In general, standardized tests are reliable (the results of what the tests measure are consistent across the student population).

SA A U D SD

2. In general, standardized tests are valid (the tests measure what they say they will measure).

SA A U D SD

3. A student's classroom achievement is more important than his/her standardized test scores.

SA A U D SD

4. Standardized tests are uniformly administered in your school system.

SA A U D SD

5. Teachers "teach" material they anticipate will be covered on standardized tests before students take these tests.

SA A U D SD

6. Standardized test score abuse (that is, scores are used for purposes other than originally intended) exists in your school system.

SA A U D SD

7. Schools should not be compared with one another on the basis of standardized test scores.

SA A U D SD

8. Teachers are provided with standardized test scores for their students in a timely manner.

SA A U D SD

9. Teachers make use of standardized test scores in a wise manner.

SA A U D SD

10. Teachers are involved in selecting standardized tests to be used in your school system.

SA A U D SD

11. Teachers' classroom performances are evaluated on the basis of their students' standardized test scores.

SA A U D SD

12. Teachers are given adequate instruction on how to administer standardized tests.

SA A U D SD

13. Parents understand standardized test scores.

SA A U D SD

14. Teachers are given adequate training in subject curriculum areas prior to administering standardized tests.

SA A U D SD

15. Scores obtained from standardized tests are useful to the classroom teacher.

SA A U D SD

16. Administrators make use of standardized test scores in a wise manner.

SA A U D SD

17. There is too much emphasis on standardized tests in our school system.

SA A U D SD

APPENDIX B
INTERVIEW GUIDE

1. Do you spend class time specifically preparing students to take standardized tests? (If yes, go to question 2. If not, go to question 5.)
2. What type of preparation do you use?
 - a. test-taking skills
 - b. emphasis on testable subject matter
 - c. relaxation techniques
 - d. knowledge drills or games
 - e. other
3. When do you start special preparation activities to help students take standardized tests?
 - a. 1 - 2 weeks before
 - b. 1 month before
 - c. over 1 month before
4. How much time do you spend in special test preparation activities?
5. Do you think normal classroom instruction adequately prepares students to take standardized tests?
6. Are there any activities or subjects that you cut from your normal instruction because of the SOL implementation?

REFERENCES

- Arnott, L. E. (1992). An analysis of elementary teachers' attitudes toward standardized achievement testing. Master's Project, University of Dayton, Dayton, OH.
- A Better Education for a Better Future. A Workshop Conducted by the Virginia Board of Education and the Superintendent of Public Instruction, Virginia PTA/PTSA Leadership Training Conference conducted at James Madison University, July 23-25, 1999.
- Birkmire, K. J. (1993). Ninth grade proficiency, is it the best measure of learning: A survey of teacher attitudes. (ERIC Document Reproduction Service No. ED 361 406)
- Brandt, R (1989). On misuse of testing: A conversation with George Madaus. Educational Leadership, 46(7), 26-29.
- Bushweller, K. (1997). Teaching to the test. American School Board Journal, 184(9), 20-25.
- Corbett, H. D., & Wilson, B. L. (1989). Statewide testing and local improvement: An oxymoron? Philadelphia, PA: Research for Better Schools, Inc.. (ERIC Document Reproduction Service No. ED 374 551)
- Gilman, D. A., & Reynolds, L. L. (1991). The side effects of statewide testing. Contemporary Education, (62)4, 273-78.
- Gordon, S. P., & Reese, M. (1997). High-stakes testing: Worth the price? Journal of School Leadership, 7(4), 345-68.
- Green, K. E., & Stager, S. F. (1986). Measuring attitudes of teachers toward testing. Measurement and Evaluation in Counseling and Development, 19(3), 141-50.
- Gronlund, N. E., & Linn, R. C. (1990). Measurement and evaluation in teaching (6th ed.). New York: Macmillan Publishing Co., Inc., 453-479.
- Haney, W., & Madaus, G. (1989). Searching for alternatives to standardized tests: Whys, whats, and whithers. Phi Delta Kappan, 70(9), 683-87.

Jaeger, R. M. (1991). Legislative perspectives on statewide testing. Phi Delta Kappan, 73, 239-242.

Madaus, G. F. (1991). The effects of important tests on students. Phi Delta Kappan, 73, 226-31.

Madaus, G. F., & Stufflebeam, D. L. (1984). Educational evaluation and accountability. American Behavioral Scientist, 27(5), 649-72.

Neill, D. M., & Medina, N. J. (1989). Standardized testing: Harmful to educational health. Phi Delta Kappan, 70(9), 688-97.

Nolen, S. B., Haladyna, T. M., & Haas, N. S. (1992). Uses and abuses of achievement test scores. Educational Measurement, Issues and Practices, 11(2), 9-15.

Resnick, D. P. (1981). Testing in America: A supportive environment. Phi Delta Kappan, 62 (9), 625-28.

Sacks, P. (1997). Standardized testing: Meritocracy's crooked yardstick. Change, 29(2), 24-31.

Salmon-Cox, L. (1981). Teachers and standardized tests: What's really happening? Phi Delta Kappan, 62(9), 631-34.

Smith, M. L. (1991a). Meanings of test preparation. American Educational Research Journal, 28(3), 521-542.

Smith, M. L. (1991b). Put to the test: The effects of external testing on teachers. Educational Researcher, 20(5), 8-11.

Soltz, D. F. (1992, April). Does it really matter what teachers think of tests? Mandated testing in two "low stakes" school districts. Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, CA. (ERIC Document Reproductive Service No. ED 344 901)

Sproull, L., & Zubrow, D. (1981). Standardized testing from the administrative perspective. Phi Delta Kappan, 62(9), 628-31.

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

Urduan, T. C., & Paris, S. G. (1994). Teachers' perceptions of standardized achievement tests. Educational Policy, 8(2), 137-156.