

2008

## Teachers' perspectives on the unintended consequences of high stakes testing

David Christopher Charles

*Louisiana State University and Agricultural and Mechanical College*

Follow this and additional works at: [https://digitalcommons.lsu.edu/gradschool\\_dissertations](https://digitalcommons.lsu.edu/gradschool_dissertations)



Part of the [Education Commons](#)

---

### Recommended Citation

Charles, David Christopher, "Teachers' perspectives on the unintended consequences of high stakes testing" (2008). *LSU Doctoral Dissertations*. 123.

[https://digitalcommons.lsu.edu/gradschool\\_dissertations/123](https://digitalcommons.lsu.edu/gradschool_dissertations/123)

This Dissertation is brought to you for free and open access by the Graduate School at LSU Digital Commons. It has been accepted for inclusion in LSU Doctoral Dissertations by an authorized graduate school editor of LSU Digital Commons. For more information, please contact [gradetd@lsu.edu](mailto:gradetd@lsu.edu).

TEACHERS' PERSPECTIVES  
ON THE UNINTENDED CONSEQUENCES  
OF HIGH STAKES TESTING

A Dissertation

Submitted to the Graduate Faculty of the  
Louisiana State University and  
Agricultural and Mechanical College  
In partial fulfillment of the  
requirements for the degree of  
Doctor of Philosophy

in

The Department of Educational Leadership, Research, & Counseling

by

David Christopher Charles  
B.S., Louisiana State University, 1987  
M.Ed., University of New Orleans, 1997  
May 2008

## DEDICATION

This study is dedicated to my wife, Colleen, and my sons, Nicky and Christopher. No man was ever blessed with a better family.

## ACKNOWLEDGMENTS

Concluding this study has definitely been an immense challenge and could not have been accomplished without the assistance and support of many individuals. I am very grateful for the understanding and support of my wife, Colleen. She has made numerous sacrifices to assist me with my work on this project. She has spent many hours reading the drafts and offering suggestions. The smartest thing I ever did was marrying her 14 years ago. This achievement is hers as much as mine.

Also, my sons, Nicky and Christopher, who have had to do without a daddy on more than one occasion due to time spent on this project. I am enormously proud of them.

Furthermore, I would like to thank my parents, Delton and Jacquelyn, who have been my role models and a constant source of encouragement. I love them dearly. Their love and support of my endeavors can never be repaid.

Also, I would like to thank my brother-in-law, Jaimie Hebert whose help was invaluable. He is a great brother-in-law and a great friend.

Dr. Charles Teddlie has been wonderful through this entire process. He has supplied excellent feedback that has helped me to develop in my research. Dr. Teddlie's understanding of research has been priceless. I have greatly appreciated his assistance and direction. I am also appreciative of the help of the rest of my committee Dr. Eugene Kennedy, Dr. Kim MacGregor, Dr. Wade Smith, Dr. Joe W Kotrlik and Dr. Earl Cheek.

Finally, I would also like to thank my family, friends, and co-workers who have all provided me with input, encouraged me to continue, and provided support in a variety of ways.

I would also like to thank all of the teachers who took the time to complete the surveys and interviews. These very special people give their all for the betterment of the children of Jefferson Parish. Your contribution is appreciated.

TABLE OF CONTENTS

ACKNOWLEDGMENTS ..... iii

LISTS OF TABLES..... vii

LISTS OF FIGURES ..... ix

ABSTRACT..... x

CHAPTER

1: INTRODUCTION ..... 1

    Statement of the Problem..... 1

    Purpose of the Study ..... 3

    Framework of the Study..... 4

    Significance of the Study ..... 5

    Research Hypotheses ..... 6

    Research Questions ..... 7

    Definition of Terms..... 7

    Delimitations and Limitations..... 10

    Summary of Chapter 1 ..... 10

CHAPTER 2: LITERATURE REVIEW ..... 12

    Introduction..... 12

    Introduction to High Stakes Testing ..... 13

    The History of High Stakes Testing in the USA..... 15

    The Role of Government and the Courts ..... 18

    Arguments For and Against High Stakes Testing..... 21

    Theory and School Improvement..... 25

    The Effects of High Stakes Testing on Classroom Practices and Students ..... 28

    Potential Effects of High Stakes Testing ..... 30

    Classroom Practices, Including Test Preparation ..... 32

    Pressure ..... 35

    Teacher Morale and Commitment to the Profession ..... 36

    A Review of the Literature on Teachers’ Perceptions of Testing Programs ..... 39

    Mixed Methods Research Design ..... 40

        Recent Developments in Mixed Methods Research..... 42

        Why Use Mixed Methods? ..... 43

    Importance of the Two Independent Variables..... 45

    Summary of Chapter 2 ..... 47

CHAPTER 3: METHODOLOGY ..... 49

    Introduction..... 49

    Research Hypotheses ..... 50

    Research Questions..... 51

    Design For The Study ..... 52

    Phase I Methodology for Study: Instrument Development ..... 53

Pilot Study .....	56
Mixed Method Data Collection Procedures .....	59
Phase II: Quantitative Phase .....	62
Determination of School Performance Score (SPS).....	62
Determination of Socioeconomic Status (SES).....	64
Survey Instrument.....	66
Mixed Methods Sampling Procedures.....	67
Administration of the Survey.....	69
Phase III: Qualitative Phase .....	69
Mixed Methods Analysis.....	70
The Mixed Method Inference Process .....	72
Researcher’s Role .....	74
IRB and Jefferson Parish Public School System Approval .....	75
Summary of Chapter 3 .....	75
CHAPTER 4: QUANTITATIVE RESULTS .....	77
Introduction.....	77
Data Collection .....	78
Descriptions of Participating Schools .....	78
1) Poor SPS Score – Lower SES .....	79
School A: Campus Description.....	79
School B: Campus Description.....	80
2) High SPS – Lower SES .....	81
School C: Campus Description.....	81
School D: Campus Description.....	81
3) Poor SPS Score – Higher SES.....	82
School E: Campus Description .....	82
School F: Campus Description .....	83
4) High SPS – Higher SES .....	84
School G: Campus Description.....	84
School H: Campus Description.....	84
Results from Phase II Study.....	85
Independent Variables in the Study.....	85
Descriptive Statistics for Independent Variables.....	87
Classroom Practice Variable.....	88
Perceived Pressure Variable .....	89
Degree of Commitment Variable.....	91
Analysis of Research Hypotheses .....	93
Rationale for Analysis .....	93
Research Question 1 .....	93
Research Question 2 .....	95
Research Question 3 .....	98
Summary of Chapter 4.....	99
CHAPTER 5: QUALITATIVE RESULTS .....	103
Introduction.....	103
Research Questions.....	104

Participants .....	105
Data Collection Procedures .....	106
Data Analysis Procedures .....	107
Instruction .....	109
Teaching to the Test.....	109
Neglecting Subjects .....	112
Time .....	113
Fairness .....	115
Focus on Instruction .....	117
Pressure .....	118
Students' Pressure.....	118
Teachers' Pressure .....	119
Commitment .....	121
Summary of Chapter 5 .....	125
CHAPTER 6: SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS .....	129
Introduction.....	129
Summary of the Study .....	129
Discussion .....	134
Implications of the Study .....	140
Recommendations for Future Research .....	143
Summary of Chapter 6 .....	144
REFERENCES .....	146
APPENDIX A: STATE BY STATE DATA CONCERNING HIGH STAKES TESTING .....	162
APPENDIX B: EFFECTS OF HIGH STAKES TESTING .....	168
APPENDIX C: SURVEY INSTRUMENT .....	173
APPENDIX D: INTERVIEW PROTOCOL .....	177
APPENDIX E: PERMISSION LETTER .....	180
APPENDIX F: INTERVIEW PERMISSION LETTER .....	182
APPENDIX G: INSTITUTIONAL REVIEW BOARD APPROVAL.....	184
APPENDIX H: JEFFERSON PARISH PUBLIC SCHOOL SYSTEM APPROVAL.....	186
VITA.....	188

## LISTS OF TABLES

Table 1.1: Levels of Corrective Actions .....	2
Table 1.2: Components of School Performance Score .....	4
Table 2.1: Phi Delta Kappan/Gallup Poll .....	22
Table 2.2: Phi Delta Kappan/Gallup Poll .....	22
Table 2.3: Potential Effects of High Stakes Testing .....	31
Table 3.1: Sources of the Questions Included in the Survey Used for this Study .....	55
Table 3.2: Rotated Factor Matrix.....	59
Table 4.1: Effective Sample Size.....	86
Table 4.2: Responses Concerning Classroom Practices .....	88
Table 4.3: Means and Standard Deviations Concerning Classroom Practices .....	89
Table 4.4: Responses Concerning Perceived Pressure .....	90
Table 4.5: Means and Standard Deviations Concerning Perceived Pressure .....	90
Table 4.6: Responses Concerning Commitment.....	91
Table 4.7: Means and Standard Deviations Concerning Commitment.....	92
Table 4.8: Tests of Between-Subjects Effects (SPS) Dependent Variable: Practice .....	95
Table 4.9: Tests of Between-Subjects Effects (SPS*SES) Dependent Variable: Practice .....	96
Table 4.10: Tests of Between-Subjects Effects (SPS) Dependent Variable: Pressure .....	97
Table 4.11: Tests of Between-Subjects Effects (SPS*SES) Dependent Variable: Pressure .....	97
Table 4.12: Tests of Between-Subjects Effects (SPS) Dependent Variable: Commit.....	98
Table 4.13: Tests of Between-Subjects Effects (SPS*SES) Dependent Variable: Commit.....	100
Table 5.1: Teachers Interviewed.....	107
Table 5.2: Hypotheses Results .....	126
Table 5.3: Research Questions Results.....	127



Table A.1: Exit Examinations.....	163
Table A.2: Current Participation In High Stakes Testing And Content Areas .....	166
Table B.1: Effects on Curriculum and Instruction.....	169
Table B.2: Effects on Student Learning.....	170
Table B.3: Effects on Attitudes and School Climate .....	172

## LISTS OF FIGURES

Figure 3.1: QUAN – QUAL Methodology for Phases II & III.....	53
Figure 3.2: Sampling Procedures Surveys .....	68
Figure 3.3: Sampling Procedures Interviews .....	69
Figure A.1: Promotion Exams.....	165
Figure A.2: State Exit Exams .....	165

## ABSTRACT

A mixed methods design was utilized that was divided into three phases to verify and explore high stakes testing's effects on teachers' perceptions regarding classroom practices, pressure, and commitment to the educational profession.

Phase I utilized previous surveys and a peer review to create a knowledge base to generate a survey instrument that measured the three areas assumed to be affected by high stakes testing (commitment, pressure, and classroom practice). The survey instrument that was created was piloted.

Throughout Phase II there was a series of three-step analysis: First, the means and standard deviations from the results of the surveys were divided into the four cells and presented. Second, one-way ANOVAs were reported (with poor or high SPS scores as the independent variables) that test each of the three hypotheses. Third, two-way ANOVAs were reported (with poor or high SPS scores and lower or higher socioeconomic status (SES) as independent variables) to assess the effect that these variables jointly have on the dependent variables.

The results of the quantitative portion of this study were that how well students performed on the high stakes testing and the SES of students at the schools had little effect on their teachers' perceptions and responses to the testing program. All three hypotheses were not confirmed. The teachers' overall scores were all above average indicating that the three areas of study were present in all situations.

During Phase III, two teachers were interviewed from each school for a total of sixteen teachers. All of the teachers interviewed stated that LEAP 21 testing did affect their instructional planning, learning strategies, and curriculum content. Such practices as teaching to the test, neglecting subjects, sequencing, and time allotment were greatly affected.

All of the teachers interviewed stated that LEAP 21 testing forced them to devote some time to test preparation. Teachers provided a range of 1/3 of class time to a 100 %.

There were many factors that were contributing to a lessening of commitment to the educational profession from some educators, especially the younger ones who have a less of a vested interest in the profession

## CHAPTER 1: INTRODUCTION

### Statement of the Problem

All schools for miles and miles around  
Must take a special test  
To see who's learning such and such - -  
To see which school's the best.  
If our small school does not do well,  
Then it will be torn down,  
And you will have to go to school  
In dreary Flobbertown (Seuss, Pelutsky, & Smith, 1998, p. 21).

Dr. Seuss wrote the book, *Hooray for Diffendoofer Day*, four years before high stakes testing in Louisiana began and ten years before its relevance was felt in Jefferson Parish. In 2005, two schools, Bunche Middle School and St. Ville Elementary were closed and their students, teachers, and administrators were sent packing to other schools. These schools were closed, largely, due to their poor test scores.

Educators perceive pressure for their students to score well on these tests from all levels – federal, state, and district. The federal government has passed “No Child Left Behind”. This piece of legislation was signed on January 8, 2002. The stated goals of this act are: to institute strong accountability standards for schools and students, expanded flexibility and local control, expanded option for parents, and an emphasis on teaching methods that have been proven to work (Goldhaber, 2002).

The accountability component has been the most controversial part of the Act (discussed in Chapter 2). It has added subgroups (minorities, special education, etc.) to the accountability system. The federal government monitors that the special groups achieve annual goals that are set and can punish those schools through the states that do not reach their goals of Adequate Yearly Progress (AYP).

The State of Louisiana has taken the cause of accountability to heart and is one of the most punitive regarding poor test scores (Johnson & Johnson, 2006). In this state, the results of the tests determine if fourth and eighth graders pass to the next grade level and in high school, whether they graduate. Also, the school and its members are held accountable. The different levels of school improvement and their consequence can be seen in Table 1.1 (LDE, 2000).

Table 1.1:

Levels of Corrective Actions

**Level I** In the first two-year growth cycle, schools with a SPS of 30 or below are placed in Level I corrective actions. These schools work with District Assistance Teams, utilizing the School Analysis Model, a state diagnostic process, to identify needs, redevelop school improvement plans, and examine use of school resources. The legislature created a School Improvement Fund to assist such schools.

**Level II** Level I schools showing inadequate growth over a two-year growth cycle enter Level II corrective actions. Assigned to a school by the LDE, a highly trained Distinguished Educator (DE) works as an advisor to help the school improve student achievement and publicly reports school improvement recommendations to the school board. Districts must publicly respond to the recommendations. Parents whose children attend a school labeled as Academically Unacceptable have the right to transfer their child (ren) to a higher performing public school in districts that are not under judicial mandates of desegregation.

**Level III** Level II schools showing inadequate growth over a two-year growth cycle enter Level III corrective actions. The DE continues as an advisor; parents continue to have an option to transfer their child (ren) to a higher performing public school in districts that are not under judicial mandates of desegregation. By spring of the first year at this level, the district must submit a Reconstitution Plan to the SBESE for approval. If the school does not show sufficient growth by the end of the first year, it must be reconstituted at the beginning of the following year, once the reconstitution plan is approved by the SBESE. If the Plan is not approved, the school then loses its State approval status and funding.

At the district level, school board members are feeling the heat. New principals of schools are often told by their school board members upon receiving the appointment that all that is needed to succeed is to keep the LEAP scores high. Other areas of importance to the overall success of a public school are rarely discussed.

The yearly results of these accountability scores have a great impact on the schools and its members. Studies regarding this topic are very relevant but relatively new. More studies are needed from those closest to the day-to-day effects of high stakes testing.

#### Purpose of the Study

Since 1999, the development of Louisiana's assessment and school accountability system has been consistent and exceeded the standards-based reform efforts taking place across the country. The accountability system in Louisiana has as its centerpiece a high stakes testing assessment known as the Louisiana Educational Assessment Program for the 21<sup>st</sup> Century (LEAP 21). LEAP 21 constitutes Louisiana's criterion-referenced testing program. These tests measure to what extent a student has mastered Louisiana's content standards. The Louisiana Department of Education posits the following:

Louisiana's high stakes testing policy is an important part of Reaching for Results, an educational reform system designed to improve student achievement. The LEAP 21 tests are designed to ensure that grade 4 and grade 8 students have adequate knowledge and skills before moving on to the next grade (LDE, 2004, p. 1).

The expectations from these reforms are that they will improve academic achievement by creating higher expectations and thereby focusing greater effort and resources on student learning. However, critics raise a variety of objections, including "the fear that higher standards without additional resources may worsen educational inequities or decrease teacher professionalism... [also] emphasis on assessments (even good ones) might narrow the curriculum and encourage teachers to teach the test "(Taylor, Shepard, Kinner, & Rosenthal, 2003, p.2).

Whether you believe that high stakes testing is right or wrong, it is obvious that teachers and their classroom practices are expected to be an important intervening variable that will determine to a large degree the effects of reforms on student learning. In order to achieve full

understanding of these intense reforms and their effect, it is important to collect the thoughts and experiences of those most closely involved (i.e., the teachers).

The purpose of this study is to assess the perceptions of Jefferson Parish teachers toward LEAP 21 Testing and how high stakes testing affected school improvement. Comparisons will be made between schools with high and low socioeconomic status and comparisons will be made between elementary schools with low SPS scores and those elementary schools that have high SPS scores. School Performance Scores (SPS) are based upon students’ test scores on LEAP and the Iowa Tests, as well as their dropout and attendance records (see Table 1.2). Special emphasis will be given to the teachers’ perceptions regarding the effects of high stakes testing on classroom practice, teacher morale/commitment, and perceived pressure.

Table 1.2:

Components of School Performance Score (SPS) Accountability Indicator Weighting Factor

LEAP 21 Tests	<b>60%</b>
iLEAP Tests	<b>30%</b>
<b>For schools without grades 7 or 8:</b> Student Attendance	<b>10%</b>

Framework of the Study

I have chosen mixed methods research to accomplish the goals this study. As Greene et al. (1989) proposed there are five functions of mixed methods: triangulation, complementarily, development, initiation, and expansion.

The first two functions of mixed methods (triangulation and complementarity) are the fact that mixed methods lead to multiple inferences that confirm or complement each other. The other three functions (development, initiation, and expansion) are more related



to mixed methods studies in which inferences made at the end of one phase (e.g., QUAL) lead to the questions and/or design of a second phase (e.g., QUAN) (Tashakkori & Teddlie, 2003, p. 16).

First, I utilized a quantitatively designed survey to create baseline data that was confirmatory in nature. Then I employed qualitatively designed interviews, which was exploratory in nature to generate a deeper understanding of the knowledge base. The former provides greater breadth, while the latter provides greater depth.

### Significance of the Study

High stakes testing has assumed a prominent role in the last decade in an effort to improve education (Hursh, 2005). “ At a cost of millions, even billions, of dollars and at the expense of valuable student, teacher, and administrator time, testing advocates and many policymakers still view testing as a significant ... tool in educational improvement” (Herman & Golan, 1991).

Previous research has indicated that high stakes testing has resulted in increased pressure on teachers, a decrease in teachers’ morale/commitment, and a change in classroom practices (Johnson & Johnson, 2002; Pedulla et al., 2003, Yeh, 2005).

The significance of the present study lies in three areas: (a) utilizing additional quantitative data to identify the actual effects of high stakes testing in the areas of morale/commitment, pressure, and classroom practices; (b) discovering if past success affects the educational communities’ perceptions toward high stakes testing; and (c) making a contribution to the literature by adding to the understanding of previous findings through qualitative data.

The first area, bringing additional quantitative to the actual effects of high stakes testing in the areas of morale/commitment, pressure, and classroom practices involved quantitative data derived from a survey given to elementary school teachers. The survey will verify or contradict findings from surveys introduced to teachers from other states with high stakes testing. The

results of this study's survey will show if teachers in Jefferson, Louisiana are consistent in their perceptions with previous research.

Concerning the second area, discovering if past success affects the educational communities' perceptions toward high stakes testing will be obtained by comparing the results of surveys that were answered by schools which had a history of past testing success and those teachers which did not have such a history. This comparison, which has not been made in previous literature, may be important to explain inconsistencies in previous research.

Finally, the third area of significance for the study is making a contribution to the literature by adding to the understanding of previous findings through qualitative data. This contribution will be the result of interviews that are informed by the surveys. The qualitative data will add depth and understanding to the current literature.

This study will look at the LEAP21 through the eyes of elementary school teachers. The teachers who deal with its impacts daily will provide insight into areas of needed further research.

This study is guided by research questions that explored high stakes testing's effects on classroom instruction, teacher morale/commitment, and perceived pressures to score high. These areas are important because of their potential to leave such lasting effects, positive or negative, on the students who take these tests, the adults who administer them, and the institutions that support both.

### Research Hypotheses

1. Jefferson Parish teachers from schools that produce high SPS scores will perceive that LEAP 21 testing has affected classroom practice more than teachers from schools that have scored poorly.

2. Jefferson Parish teachers from schools which produce low SPS scores will perceive that LEAP 21 testing has created pressure to spend more time on test preparation (teaching to the test) than teachers from schools that have achieved high SPS scores.
3. Teachers from schools with high SPS scores will indicate that they have a higher degree of commitment to the education profession than teachers from schools with low SPS scores.

### Research Questions

The preceding research hypotheses are confirmatory statements that will be tested using results from the quantitative component of a survey. The questions that I present now will be informed by the results from the qualitative component of the study.

1. How does test preparation (teaching the test) affect teachers' instructional planning, learning strategies, and curriculum content and to what extent?
2. How much time do teachers perceive that students spend on test preparation and how does that amount of time compare to the time spent on instruction?
3. What effect does testing have on an educators' sense of professionalism and pride in their work? How does high stakes testing affect motivation in general?

### Definition of Terms

The following terms and operational definitions were used throughout this study:

- **Adequate Yearly Progress (AYP)**
- This is the minimum level of achievement or improvement that a school must achieve within a set time frame. The No Child Left Behind Act of 2001 (NCLB) requires that every state form its own definition of AYP. Louisiana evaluates whether schools make AYP for two components:

- SPS Component – to make AYP a school must have a baseline SPS of 45 or above; and
- Subgroup Component – to make AYP a school must meet requirements in test participation, and the additional academic indicator (attendance rate or non-dropout rate) (Louisiana State Education Progress Report 2004-2005).

- **Free and Reduced Price Lunch Program**

The percentage of students eligible for this federally subsidized program used as an indicator of family economic condition. Based on the U.S. Government's 2002-2003 guidelines, the maximum family income for eligibility in the Free Lunch Program is 130% of the federal poverty level, or \$23,530 annually for a family of four. The family income limit for eligibility in the Reduced Lunch Program is 185% of the federal poverty level, or \$33,485 annually for a family of four (Louisiana State Education Progress Report 2004-2005).

- **Grade-Level Expectations**

Further define the content standards and benchmarks for English language arts, mathematics, science, and social studies in grades PreK through 12 and is a statement that indicates what all students should be able to do at the end of a grade level (Louisiana State Education Progress Report 2004-2005).

- **Growth Target**

This is the amount of progress that a school needs to make to remain on track for reaching the state's SPS goal of 120.0 for 2014 (Louisiana State Education Progress Report 2004-2005).

- **High Stakes Testing**

Describes tests that have high stakes for individual students, such as grade promotion or a standard high school diploma (Cortiella, 2004).

- **iLEAP**

The Iowa Tests (NRT) augmented with criterion-referenced test items that are Louisiana specific and measure grade-level expectations that are not measured by The Iowa Tests. This assessment plan of combining the NRT and CRT is being referred to as the iLEAP or integrated Louisiana Educational Assessment Program (Louisiana State Education Progress Report 2004-2005).

- **LEAP 21**

Tests that measure how well students master the state's content standards and are administered to students in the 4th and 8th grades (Louisiana State Education Progress Report 2004-2005).

- **School Improvement (SI)**

Formerly called Corrective Actions, has six levels, five of which were applicable with the 2003-2004 Accountability release. Schools enter or move further into SI if they do not meet performance and growth requirements. These schools receive support and assistance based on their SI level. A detailed description of the rules and regulations which apply to School Improvement, are found in Bulletin 111: Louisiana School, District, and State Accountability Policy, which can be found on the Louisiana Department of Education's website at [www.louisianaschools.net/lde/bese/home.html](http://www.louisianaschools.net/lde/bese/home.html).

- **School Performance Score (SPS)**

This is the primary measure of a school's overall performance (Louisiana State Education Progress Report 2004-2005).

- **Pragmatism**

This is a deconstructive paradigm that debunks concepts such as "truth" and "reality" and focuses instead on "what works" as the truth regarding the research questions under investigation. Pragmatism rejects the either/or choices associated with the paradigm wars,

advocates for the use of mixed methods in research, and acknowledges that the values of the researcher play a large role in interpretation of results (Tashakkori & Teddlie, 2003, p.713).

#### Delimitations and Limitations

Delimitations of this study include the use of SPS scores to categorize schools as a whole and not breaking down the schools into subgroups and subsets.

This study is also delimited to public schools and not to private schools that are not involved with high stakes testing. Furthermore, magnet and alternative schools were excluded to enhance comparability. These schools often deal with a different set of variables that impact testing (Thomas, 2005).

Finally, the use of Jefferson Parish limits the study. The stringent testing guidelines of Louisiana may limit generalization to other states. Also, Jefferson Parish's unique setting of recovering from Hurricane Katrina may not generalize to other parishes.

#### Summary of Chapter 1

In this chapter the consequences of high stakes testing were introduced. Also, the purpose of this study which is to assess the perceptions of Jefferson Parish teachers toward LEAP 21 Testing and how high stakes testing affects school improvement was discussed. To achieve this purpose, mixed methods were utilized as the framework for this study.

The significance of the present study lies in three areas: (a) generating additional mixed methods data to further understand the effects of high stakes testing in the areas of morale/commitment, pressure, and classroom practices; (b) discovering if past success affects the educational communities' perceptions toward high stakes testing; and (c) making a contribution to the literature by adding to the understanding of previous findings through qualitative data.

The research hypotheses and questions that guided this study were introduced and definitions to terms used throughout the study were provided. Finally, limitations to the study were addressed.

## CHAPTER 2: LITERATURE REVIEW

### Introduction

High stakes testing as a focus of research is relatively new. However, given its profound effect on the lives of students, teachers, and administrators and its wide spread implementation, a great deal of material has been written in a short time. “Over the last 15 years, the movement for higher standards and accountability in our schools has led several states – and now the federal government with the ‘No Child Left Behind’ (NCLB) Act – to adopt test-based accountability policies” (Goldberg, 2004, p.8).

In this review, high stakes testing will be fully defined and discussed. The discussion will include the history of high stakes testing and the role of the government and the courts in its progression. Also, the discussion will include testing’s current and potential effects.

The research design employed by this study is a mixed methods design. This literature review includes an overview of teachers’ perceptions concerning testing and the utilization of mixed methodologies research. Also, in this chapter, I will discuss why it was important to utilize mixed methods for this study. Furthermore, the independent variables (socioeconomic and past test success) and the dependent variables (classroom practices, pressure, and teacher morale/commitment) employed in this study are discussed.

The research strategies employed in this study to identify the relevant literature related to high stakes testing included a computer search conducted through *Education Resources Information Center (ERIC)*, *Google Scholar*, and *Dissertation Abstracts International*. Also, a manual search of bibliographies of selected books, articles, and papers was conducted. This search generated more than 200 citations from journal articles, papers, and books that are included in the reference section of this study.



This review of the literature is divided into the areas concerning the subject and methodologies employed.

1. a review of literature concerning high stakes testing
2. a review of the effects of high stakes testing
3. a review of the dependent variables employed in this study
4. an overview of research on teachers' perceptions regarding high stakes testing
5. a review of mixed methodologies
6. a review of the independent variables employed in this study.

### Introduction to High Stakes Testing

The debate regarding high stakes testing has been very public taking place in the press and on the campuses. Although there is much disagreement, there are some points on which both sides can agree. These points are that the debate is highly emotional, the stakes are high, and that it is an issue that is in the forefront of most K-12 educators' minds. In this literature review, I plan to provide a better understanding of high stakes testing and its effects on school improvement.

For the purposes of this study, I will use Dale DeCesare's (2002) definition of high stakes testing as

a term that is used to describe programs designed to measure not only the achievement of students, but also of teachers, principals, and schools. 'High-stakes' also is used to describe assessment tools that can have a variety of consequences. In the case of schools, such consequences can range from a letter of warning for low performance to mandatory reconstitution (p.10).

Furthermore, high stakes testing generates assessments whose results have important consequences for students. Such stakes may include promotion, certification, and graduation.

Madaus (1988) considered a test high stakes if the results of the test have perceived or real consequences for students, staff, or schools.

Determining the exact number of states involved in using high stakes testing programs is more complex than first anticipated. This is due to the use of differing terms such as “accountability programs” and “efficiency programs”. Also, the states have employed a diverse set of programs that employ accountability to various degrees. Currently, twenty-six states apply some sort of consequences to their standardized testing programs. Seventeen states, including Louisiana, go so far as to withhold a child’s diploma (information retrieved from the Boston College Law School website on December 15, 2006).

Louisiana is in the minority in that it has embraced one of the strictest standards and degree of consequences in its high stakes testing. Louisiana not only ties high stakes testing to promotion and graduation, but also ties the State’s curriculum standards to these test. Current trends are that other states will develop stricter standards, especially due to the No Child Left Behind (NCLB) Act. A comparison of the Louisiana accountability to the other states is located in Appendix A.

More and more, states and school boards are using standardized test scores in order to judge schools and allocate resources. Rewards and punishments are increasingly being tied to the results of high stakes testing.

In October 1996, Chicago put 109 schools on academic probation. According to Hendrie (1996) scores from nationally normed standardized tests were a chief factor in determining who would be placed on probation. Manzo (1996) reported that Philadelphia was planning to link teacher raises and cash awards to schools based on student test scores, attendance, and graduation rates. For schools with chronically low-performing students, schools could be forced to replace up to three-fourths of their staffs (Langenfeld, Thurlow, and Scott, 1997, p. 2).

NCLB is located in the Title I section of the Elementary and Secondary Education Act. According to this legislation each state is required to initiate content and performance standards in English (reading) and math, with assessments based on them, and to add science later. These standards are to include four levels (advanced, proficient, basic and below basic). Currently,

every state but Iowa now has standards and at least some state mandated assessments.

NCLB further mandates that

[b]y the 2005 – 2006 academic year, states must assess each child every year in grades 3-8 and once during high school in math and reading/language arts based on the content and achievement standards...Until 2005 – 2006, annual testing in reading and math once in three grade spans (3 – 6, 6 – 9, 10 – 12) is required. By 2007-2008, states must add an annual science assessment in the three grade spans. Commercial norm-referenced tests will be allowed if items are added to ensure the tests cover the state standards. State assessment systems that are a mix of state exams and local assessments are also allowed – Nebraska, Maine, Rhode Island, and Vermont did this. These assessments will be the ‘primary’ method of determining progress toward the goal of all students reaching the ‘proficient’ level by 2014 (retrieved from FairTest.org on December 15, 2006).

The future looks bright for advocates of high stakes testing. To better understand this phenomenon we need to look at its past.

#### The History of High Stakes Testing in the USA

This modern reform finds its origins in Russia’s launch of Sputnik, the first successful man made space-orbiting satellite in 1957. This event at the height of the cold war resulted in an increased emphasis on education by the federal government by passing the National Defense Education Act, which provided increased funding in the area of math and science (Ravitch, 2000). This was interrupted by the civil rights movement but gained steam again when international comparisons of students showed the United States slipping. As Bunting (1999) observed, each new fix became the source of a new problem.

In 1965, Title I was passed as part of the Elementary and Secondary Education Act which was part of Lyndon Johnson’s “War on Poverty”. Gary Natriello and Edward L. McDill (1999) noted that

...the underlying premise of Title I regulations implied that schools as organizations were not important: Title I service delivery was predicated on the assumption that local compliance with federal mandates was sufficient to secure educational results for precisely those students whom the schools had the most difficulty educating. Assessment and evaluation focused on compliance with procedural requirements that were often labyrinthine. In order to comply with federal regulations, compensatory students were

segregated from others. The resulting separation between students identified as disadvantaged and low achieving from the rest simply exacerbated the isolation of Title I students and services (p. 3).

In the 1980's, federal reform initiatives in high stakes testing began to take shape. The National Commission on Excellence in Education was formed by the Secretary of Education. The commission's report was called a Nation at Risk: The Imperative for Educational Reform and was the start of an educational reform movement. This report led to different waves of reform (Smith and O'Day, 1993).

The first wave under President Reagan utilized top-down mandates for change in areas such as curriculum and graduation requirements. During this wave every state developed their own plans, which emphasized improving existing programs. However, little thought was given to changing deep seeded structures such as textbook reliance and curriculum tracks (Wallace & Graves, 1995).

Also, during the first wave of reform, the U. S. Supreme Court ruled on a court case from Florida. In this case, Debra P. v. Turlington (1981), a standard was set for how a fair opportunity to learn was legally defined. In this case the plaintiffs challenged the use of minimum competency tests as a prerequisite in order to graduate. The Supreme Court ruled that since the test measured skills that were consistent with the curriculum, the students had a fair opportunity to learn (Heubert & Hauser, 1999).

The case of Debra P. offers an especially clear illustration of a crucial distinction between appropriate and inappropriate test use. Is it ever appropriate to test students on material they have not been taught? Yes, if the test is used to find out whether the schools are doing their job. But if that same test were used to hold students 'accountable' for the failure of the schools, most testing professionals would find such use inappropriate. It is not the test itself that is the culprit...results from a test that is valid for one purpose can be used improperly for other purposes (Heubert & Hauser, 1999, p. 21).

The second wave of reform that took place in the 1990's and moved to a more bottoms-up emphasis with a focus on decentralization. Also, strides were taken toward school-based

management. During this wave accusations were made that teachers were only teaching to the test and test taking skills such as how to take multiple-choice tests (Hamilton & Koretz, 2002).

Also, during this time period, Title I was reformed in the Hawkins-Stafford amendments, first in 1988 and then subsequent revisions were made in 1994. These amendments were part of the Improving American Schools Act. “One of the most important of these changes permits schools with high concentrations of poverty students to use Title I funds school-wide, rather than only for eligible students” (Natriello & McDill, 1999, p. 4).

Within these amendments to Title I there were four provisions that highlighted academic effectiveness.

1. Improved coordination between Title I and the regular school curriculum by developing more integrated school wide approaches for meeting the needs of all students.
2. Parental involvement – The legislation specified procedures for more systematically involving parents in the planning, review, and implementation of the program through the use of written district policies.
3. School wide projects – Congress eased restrictions on the development of whole school reforms where the poverty level was 75% or greater...
4. Accountability for school performance – Congress increased its demands for program effectiveness by requiring school districts to identify schools that failed to demonstrate academic progress and then aid these institutions in developing and implementing improvement plans (Natriello & McDill, 1999, p. 9).

In President Clinton’s 1997 State of the Union address, the President implored the country to undertake

...a national crusade for education standards – not federal government standards, but representing what all our students must know to succeed in the knowledge economy of the twenty-first century...Every state should adapt high national standards, and by 1999, every state should test every fourth-grader in reading and every eighth-grader in math to make sure these standards are met...Good tests will show us who needs help, what changes in teaching to make, and which schools need to improve. They can help us to end social promotion. For no child should move from grade school to junior high, or junior high to high school until he or she is ready (Heubert & Hauser, 1999, p. 14).

The strategy to implement these goals was initiated prior to the President’s speech and was called America 2000, which hoped to raise academic achievement for all students and set

target graduation rates (Ravitch, 2000). Although President Clinton's call for testing was voluntary, by 1995, eighteen states had a prerequisite of an exit test requirement for high school graduation (Bond & King, 1995). America 2000 became Goals 2000 and each state was given the task of developing content standards (Natriello & McDill, 1999).

President Clinton was followed by President Bush in 2001. Some of the same goals stated in Goals 2000 were adopted by the Bush administration; however the focus on assessment has become much stronger. Through the previously mentioned No Child Left Behind Act, the federal government placed stricter guidelines on accountability practices (Kiely & Henry, 2001; Smith, 2005).

### The Role of Government and the Courts

Since its inception, the No Child Left Behind Act has fended off numerous court challenges. "The Supreme Court has held that Section 504 does not require 'an educational institution to lower or effect substantial modifications of standards to accommodate a handicapped person'. In fact, as is the case with the Equal Protection Clause, suits under Sections 504 challenging the applicability of exit exams to students with disabilities have not met with much success" (O'Neill, 2003, p. 648). Lawsuits involving Individuals with Disabilities Education Improvement Act (IDEA) have met with a similar fate.

In Louisiana, the Louisiana Educational Assessment Program for the 21<sup>st</sup> Century (LEAP 21) was challenged in court. In the Parents Against Testing Before Teaching v. Orleans Parish School Board, 273 F. 3d 1107 (5<sup>th</sup> Cir. 2001), the Supreme Court refused to hear its appeal in March of 2002. In this case the plaintiffs were trying to get the LEAP test thrown out.

The plaintiffs, a group of parents, challenged the overall fairness of the test and sought to bar the state and school districts from denying promotion to fourth and eighth grade students who fail it. According to plaintiffs, forty-two percent of the New Orleans districts' fourth graders and fifty-three percent of its eighth graders scored 'unsatisfactory' on the 1999 tests. The denial of certiorari lets stand the district court's

1999 ruling, which was affirmed by the Fifth Circuit, holding that, while courts have recognized a property interest in receiving a diploma, ‘no court has ever recognized a property interest in promotion’ (O’Neill, 2003, p. 654).

The Louisiana legislature reacted to the different federal programs by developing one of their own. In 1993, Louisiana Systemic Initiatives Program (LaSIP) was the first of Louisiana programs to be initiated in reaction to these federal reforms covered in this literature review (Finley, 1999). The content standards became linked to performance standards and were used to understand how well students met the standards. Before this change, Louisiana tied assessments to competencies in accordance with Act 750 of 1979.

Under Mike Foster (1996-2004), Louisiana embraced the federal reform efforts. “In 1997, the Louisiana legislature passed an act creating the School and District Accountability Commission and assigning it to ensure measures of student performance were in place. Hence, students, schools, and districts [became] accountable for student performance ...” (Mancuso, 2004, p. 33).

This led to the creation of LEAP 21, which made Louisiana the first state in the United States to require fourth and eighth graders to earn a certain score on a standardized high stakes test in order to be promoted to the next grade.

Although some states have high school exit exams that students must pass to graduate, Louisiana appears to be the first state to have in place an accountability system for earlier grades that makes passing a certain test the maximum benchmark for advancement to the next grade. Individual districts, including the Chicago school system, have policies that hold back students based on an assessment. ‘That’s the first state we know of,’ said Matthew Gondal, the vice president of Achieve Inc., a nonprofit Cambridge, Mass. group formed by state and business leaders to help promote improved student achievement. ‘I’m sure people are going to be watching closely outside of the state ...’ (Robelen, 2000, p. 25).

The No Child Left Behind Act had four areas that gained much attention. “H. R. 1 asks states to put a highly-qualified teacher in every school classroom by 2005” (retrieved from ED.gov on October 19, 2006).

A key feature of the act is its focus on highly qualified teachers. Beginning the school year [2001-2002], new teachers in schools receiving funds under the law must meet state standards as being highly qualified for the positions. Those teachers already in classrooms have 4 years to meet their state's standards" (Rose, 2002, p. 322).

According to proponents this is one of the strengths of this Act because, "[s]tudies that seek to identify the factors that improve school performance all agree that teacher quality is the critical element of success" (Sclafani, 2002, p. 43). Furthermore,

Sanders and Horn report results from the Tennessee Value – Added Assessment System, a 'massive, longitudinally merged database linking students and student outcomes to the schools and systems in which they are enrolled and to the teachers to whom they are assigned. Results show that race, socioeconomic level, and class size are 'poor predictors of student academic growth' and that the major determinant of academic growth is the quality of the teacher" (Strahan, 2003, p. 298).

However, these researchers have not allowed their data to be reanalyzed.

Another aspect of the No Child Left Behind Act is research; state academic programs must be based on scientifically validated practice. Proponents advocated "research that works. Although that sounds simple and obvious, the reality is that we have not done it" (Sclafani, 2002, p. 44). Susan Sclafani also believes this research requirement should lead to more hands-on activities that will lead to success.

A third important aspect of this Act is in the area of parental choice. Parents have the right to transfer their child out of a school which is repeatedly labeled low performing (Hombo, 2003).

Ed.gov (2005) found the following:

H. R. 1 creates meaningful options for parents whose children are trapped in failing schools and makes these options available immediately:

- Public School Choice: Parents with children in failing schools would be allowed to transfer their child to a better-performing public or charter school immediately after a school is identified as failing.
- Supplemental Services: Federal Title I funds (approximated \$500 to \$1000 per child) can be used to provide supplemental educational services – including tutoring, after school services, and summer school programs – for children in failing schools.
- Charter Schools: H. R. 1 expands federal support for charter schools by giving parents, educators and interested community leaders greater opportunities to create new charter schools (retrieved on October 19, 2006).



Finally, as discussed earlier, the most controversial aspect of the No Child Left Behind Act relates to accountability (high stakes testing). In this Act, it is left to the state to set student achievement standards and to create assessments which align with these standards (Sclafani, 2002). These standards must be at least equivalent to the standards set in the National Assessment of Educational Progress (NAEP) (Hombo, 2003).

Essentially, states must create an accountability system that includes all students. Progress in mathematics, reading, and science must be measured yearly. Schools, which do not demonstrate this progress over two years, must develop corrective action plans. If these plans do not produce results, schools may face changes in staffing and curriculum, or a possible state takeover. While schools receiving Title I funds have long been required to conduct assessments, such assessments were required only in one grade per span. Under No Child Left Behind, every child must be tested yearly in grades 3 through 8 in reading and mathematics (by the 2005-2006 school year) and in science (by the 2007-2008 school year)” (Kymes, 2004, p. 4).

Furthermore,

“...states must develop separate progress goals for subgroups of students, including economically disadvantaged students, students from major ethnic and racial groups, students with disabilities, and limited English proficiency students, as well as all public school students” (Goertz & Duffy, 2003, p. 7).

#### Arguments For and Against High Stakes Testing

In the debate over whether high stakes testing has positive or negative effects of school improvement and school practices, one side is dominant over the other. Far more articles and books have been written on the behalf of those that oppose the use of high states testing than proponents for them. High stakes testing is advocated, however, by a majority of the parents and the former Secretary of Education Ron Paige and current Secretary Margaret Spellings. The Phi Delta Kappan/Gallup poll consistently measures support from parents since 1978 (Heubert & Hauser, 1999). The public believes that the amount of achievement testing in schools is just about right, and a majority of respondents support additional testing. The 40% of parents

that say there is about the right amount of testing and the 17% saying there is not enough constitute a majority in support of testing. Two of the questions are provided in

Table 2.1 and 2.2.

Table 2.1:

Phi Delta Kappan/Gallup Poll

In your opinion, is there too much emphasis on achievement testing in the public schools in your community, not enough emphasis on testing, or about the right amount?

	National Totals					No Children In School					Public School Parents				
	'05 %	'04 %	'02 %	'01 %	'00 %	'05 %	'04 %	'02 %	'01 %	'00 %	'05 %	'04 %	'02 %	'01 %	'00 %
Too much	36	32	31	31	30	35	30	30	29	28	39	36	32	36	34
Not enough	17	22	19	22	23	17	23	20	22	26	17	20	14	20	19
About the right amount	40	40	47	44	43	39	40	46	45	41	43	43	54	43	46
Don't know	7	6	3	3	4	9	7	4	4	5	1	1	a	1	1

<sup>a</sup>Less than one-half of 1%.

Table 2.2:

Phi Delta Kappan/Gallup Poll

In your opinion, should one of the measurements of a teacher's quality be based on how well his or her students perform on standardized tests or not?

	National Totals		No Children In School		Public School Parents	
	'05 %	'04 %	'05 %	'04 %	'05 %	'04 %
Yes, should	52	49	53	50	52	49
No, should not	44	47	43	45	46	49
Don't know	4	4	4	5	2	2

Locally, the Public Affairs Research Council (2002) conducted a poll in 2001. Their findings were:

Although vocal, the opposition is mainly centered in the New Orleans area and lacks strong support. A recent poll of nearly 500 registered Orleans Parish voters found that 77% either 'strongly approved' or 'approved' of the current LEAP testing in schools. There was also support across racial lines. Some 93% of white voters polled and 63% of black voters polled indicated approval of LEAP testing (p. 9).

Also, understandably, the educational testing services (Achieve, Inc., Educational Testing Service, The College Board, Kaplan and the Association of Test Publishers) have sponsored writings that advocate their position. Furthermore, Richard Phelps (2003) wrote a scathing book attacking those that opposed high stakes testing. Phelps argues that much of the research conducted by education insiders concerning high stakes testing is based on ideological preference or profound self-interest. He believes that it should not be surprising that these educators arrive at emphatically anti-testing conclusions. He notes that external and high stakes testing in particular attracts a cornucopia of invective. Much, if not most, of this hostile research, according to Phelps, is passed on to the public by journalists as if it were neutral, objective, and independent.

Finally, surprisingly, the American Psychology Association (APA) (Carpenter, 2001) released a supporting position; however, they found high stakes tests acceptable in very narrow circumstances.

Their positions are well summarized by Amrein & Berliner, although they were opponents of high stakes testing.

Proponents argue that:

- students and teachers need high stakes tests to know what is important to learn and to teach;
- teachers need to be held accountable through high stakes tests to motivate them to teach better, particularly to push the laziest ones to work harder;
- students work harder and learn more when they have to take high stakes tests;

- students will be motivated to do their best and score well on high stakes tests; and that
- scoring well on the test will lead to perceptions of success, while doing poorly on such tests will lead to increased effort to learn.

Supporters of high stakes testing also assume that the tests:

- are good measures of the curricula that is taught to students in our schools;
- provide a kind of ‘level playing field’, an equal opportunity for all students to demonstrate their knowledge;...

Finally, the supporters believe that:

- teachers use test results to help provide better instruction for individual students;
- administrators use the test results to improve student learning and design better professional development for teachers... (Amrein & Berliner, 2002, p. 4).

Amrein and Berliner then went on to say that all of these assertions have been researched both quantitatively and qualitatively along with interviews of those that work or participate in high stakes testing environments.

A reasonable conclusion from this extensive corpus of work is that these statements are true only some of the time, or for only a modest percent of the individuals who were studied. The research suggests, therefore, that all of these statements are likely to be false a good deal of the time. And in fact, some research studies show exactly the opposite of the effects anticipated by supporters of high stakes testing” (Amrein & Berliner, 2002, p. 5).

Much of the debate against the test considers its political context and fairness.

Virtually all relevant experts and organizations condemn the practice of basing important decisions, such as graduation or promotion, on the results of a single test. The National Research Council takes this position, as do most other professional groups (such as the American Education Research Association and the American Psychological Association), the generally pro-testing American Federation of Teachers, and even the companies that manufacture and sell the exams” (Kohn, 2000, p. 61).

Kohn also states that evidence shows that many teachers are teaching to the test. McNeil (2000) found in his research that school reform efforts that centered on testing, greatly distorted the educational experiences of students in urban schools. She found that as schools focused more and more on test preparation and teaching to the test, test scores increased, meanwhile the quality of teaching and learning was both compromised and depreciated (Wright, 2002, p. 4).

Mancuso (2004) states that

Hauser, Pager, and Simmons (2000) suggest that differences in retention rates of Black and White students can largely be explained by social and economic factors. However, differences in test scores are generally larger than what would be expected from social and economic differences. The difference suggests that tying test scores to promotion purposes has a disparate impact on racial and ethnic minority students (Hauser et al, 2000) (p. 39).

Klein, Hamilton, McCaffrey, & Stecher (2000) also support these views. Furthermore, some researchers believe that society is too pluralistic and multicultural to lend itself to one important test (Strike, 1998).

Politically, McDonnell, McLaughlin and Morrison (1997) noted that standards-based reform has mobilized diverse ideological interest groups...they caution that ‘to talk about the institutional arrangements assumed in the standards-based policy framework is to pose a question about who has authority to define and implement standards and to ask whether consensus is possible among all these different interests’ (p. 32)” (Wright, 2002, p. 6).

### Theory and School Improvement

Social Cognitive Theory (Bandura, 1986) provides a framework for predicting and understanding both individual and collective behavior. Also, it explains how that behavior is developed and supported.

In the model, the interaction between the person and behavior involves the influences of a person’s thoughts and actions. The interaction between the person and the environment involves human beliefs and cognitive competencies that are developed and modified by social influences and structures within the environment. The third interaction, between the environment and behavior, involves a person’s behavior determining the aspects of their environment and in turn their behavior is modified by that environment (Davis, 2006, p.1).

According to the theory there are three factors environment, people, and behavior. These elements are constantly influencing each other. Behavior is not only due to the environment and the individual; likewise the environment is not only the result of the person and behavior (Glanz et al, 2002). It is important to understand that individuals (such as teachers) are both contributors and products of their organizations (schools).

Getzels and Guba (1957) presented social systems theory that explains organizational behavior in terms of how it addresses the social needs of its members. This theory highlights and provides a framework for comprehending the complex nature of social systems that exists in schools. Getzels and Guba (1957) also note that performance is the result of interplay between an individual's personality and the same person's role in the organization.

Usefulness of the theory lies in the interdependent dynamic nature of the process of education implied in the ideographic (e.g., human personality) and nomothetic (e.g., individual goals, group goals, and expectations) dimensions discussed. The theory supports research that examines the functions and processes associated with organizational structure, goals, culture, political influences, and individual needs within the education system. Important to this framework is the significance of the interconnected dynamic nature of the education organization (Clark, 2005, p. 30).

The theories presented in this section support a knowledge base for understanding that educational institutions are complex and dynamic organizations or social systems. Nikki Clark (2005, p. 31) states:

This view of schools means that systemic reform is dependent upon capacity at multiple levels within the education bureaucracy. For example, research on systemic reform has shown that (a) shared vision for reform, (b) instructional guidance for the realization of the vision, (c) adequate resources, efficient delivery of services, and accountability are necessary components for systemic reform that results in improved student achievement (Goertz, Floden & O'Day, 1995). Absence of the capacity to achieve any one of the components has the potential to impact school effectiveness in a negative manner... Research has also shown that school effectiveness was directly impacted by a number of factors including school culture, teacher self-efficacy, and leadership. External factors such as accountability, in addition to policy-guided school improvement efforts from federal, state, or district sources also have a direct impact on school effectiveness.

Any changes, especially external changes, exerted on organizations (schools) that are so dynamic, complex, and open requires careful decisions based on informed research (Hoy & Miskel, 1996). High Stakes Testing has had a profound effect on schools in Jefferson Parish. This research will add to the knowledge base of what are those effects.

Amrein & Berliner used Heisenberg's Uncertainty Principle to illustrate that high stakes testing greatly affects these social systems.

For many years the research and policy community has accepted a social science version of Heisenberg's Uncertainty Principle. That principle is the more important that any quantitative social indicator becomes in social decision-making, the more likely it will be to distort and corrupt the social process it is intended to monitor. When applied to a high-stakes testing environment, this principle warns us that attaching serious personal and educational consequences to performance on tests for schools, administrators, teachers, and students, may have distorting and corrupting effects. The distortions and corruptions that accompany high-stakes tests make inferences about the meanings of the scores on those tests uncertain. If there is uncertainty about the meaning of a test score, the test may not be valid (2002, p. 3).

On a more individual level, Self-determination theory (SDT) is a theory concerning individual motivation. This theory addresses the progression and functioning of a person's personality within social contexts. The theory begins with the assumption that people are born with a propensity toward emotional and mental growth. They desire to master challenges and to incorporate their experiences into a logical sense of who they are. This process is not automatic, according to the theory, but requires continued nurturing and support from the individual's social environment in order to work successfully. The social environment can either sustain or frustrate this process.

According to the Self-determination theory (SDT) the effects of testing on an individual's motivation depends on the meaning that those involved give to the event. This theory stipulates that the meaning of testing can be informational, controlling, and less than motivating (Deci & Ryan, 1985).

Evaluations and assessments have informational significance when they provide relevant feedback in a relatively supportive way. That is, when an assessment provides individuals with specific feedback that points the way to being more effective in meeting challenges or becoming more competent, and does so without pressuring or controlling the individuals, it tends to have a positive effect on self-motivation. Evaluations and assessments have controlling significance, in contrast, when they are experienced by the individuals as pressure toward specified outcomes or when they represent a means by which the evaluators attempt to control the activity and effort of the individuals or units being tested. According to SDT, when evaluations have controlling significance they tend to produce compliance and rote memorization, but they ultimately undermine self-motivation, investment, and commitment in the domain of activity being evaluated (Deci & Ryan, 2007, p. 2).

Finally, according to SDT, high stakes testing can hurt motivation when the tests convey uselessness or injustice to those involved. When the LEAP tests are perceived to be beyond the reach of some of the students being tested, it can damage all motivation and lead to an abandonment of effort.

### The Effects of High Stakes Testing on Classroom Practices and Students

Consensus has always been elusive among educational programs. The most important aspect to be discussed is how high stakes testing effect classroom practices and students individually. The research has shown that high stakes testing does impact how educators teach.

For instance, as a result of the testing in North Carolina, 59 percent of elementary, middle, and high school teachers reported changing their teaching methods (Yarbrough, 1999). In another study of one North Carolina County, 74 percent of teachers reported changing their methods in writing, 52 percent in math, and 48 percent in reading (Jones and Johnston, 2002). Similarly, Barksdale-Ladd and Thomas (2000) found that 75 percent of teachers in two other large states changed their instructional practices in response to high stakes testing (Jones, Jones & Hargrove, 2003, p. 37).

These changes in instructional practices suggest that many teachers are trying to adapt their teaching to meet the increasing demands of high stakes testing. Another way in which class practices have been affected is that student – centered practices have been replaced by teacher-centered practices.

To be sure, many city schools that serve low-income children of color were second rate to begin with. Now, however, some of these schools in Chicago, Houston, Baltimore, and elsewhere, are arguably becoming third rate as the pressures of high stakes testing lead to a more systematic use of low-level, drill-and-skill teaching, often in the context of packaged programs purchased by school districts (Kohn, 2000, p. 325).

Surveys show that teachers perceive that high stakes tests hamper creativity. One teacher reported that

I'm not the teacher I used to be. I used to be great, and I couldn't wait to get to school every day because I loved being great at what I do. All of the most powerful teaching tools I used to use every day are no good to me now because they don't help children get ready for the test, and it makes me like a robot instead of a teacher (Barksdale-Ladd and Thomas, 2000, p. 392).



Some research has found that curriculum has been negatively impacted.

Smith and Rottenberg (1991) reported on an extensive research study... They concluded among other things that (1) external testing reduces the time available for ordinary instruction, (2) testing affects what elementary schools teach – in high stakes environments, schools neglect material that external tests exclude, (3) external testing encourages use of instructional methods that resemble tests, and (4) ‘as teachers take more time for test preparation and align instruction more closely with content and format, they diminish the range of instructional goals and activities’ (p. 11) (Mehrens, 1998, p. 9).

Smith, Edelsky, Draper, Rottenberg, & Cherland (1998) also found that curriculum was adversely affected. In contrast to the previously cited research,

a survey of North Carolina teachers found that since the implementation of high stakes testing, 26.8 percent of elementary teachers reported using more student-centered instruction, compared to only 12.1 percent who reported using less student-centered instruction (Jones et al., 1999)... [Furthermore], Firestone et al. (2001) studied fourth grade math and science teachers’ instruction in New Jersey and found that ‘teaching to the test is encouraging teachers to consider more inquiry-oriented instructional practice’ (p. 11)” (Jones, Jones, & Hargrove, 2003, p. 44).

Improvements in teaching methods have also been identified in a research study of two Kentucky teachers’ math teaching methods. In response to Kentucky’s assessment program, one school’s curriculum committee changed their mathematics curriculum. The result was that one schoolteacher claimed that they had “a much tighter more comprehensive math program” (Borko & Elliot, 1999, p. 396).

Many researchers focused on high stakes testing’s effects on teaching practices, others, however, have focused on the student individually. Such is the case in the area of grade retention, students who have passed their teachers but failed a single test and thus are being forced to repeat a grade has been fodder for the critics. Darling-Hammond and Falk (1997) call grade retention “a crude concept that derives from the assembly line model of schooling...[and that] ignores questions of whether the child was appropriately taught the first time, whether doing the same thing over again is likely to be successful, and whether the educational environment itself, rather than the child is lacking” (p. 191). Also, those who oppose high stakes

testing imposed retention argue that retention negatively impacts dropout rates and increases negative academic self-esteem (Roderick, Jacob, & Bryk, 2002). Jones, Jones, & Hargrove (2003) point out that when schools retain students just one time, the dropout rate of these students rise 50 percent and two or more grade level retentions push the chance of dropping out of school to 90 percent. Furthermore, researchers have made the argument that high stakes testing leads to lower rates of promotion and graduation of students (Heubert & Hauser, 1999; Hochschild & Scott, 1998).

Louisiana's retention rates were studied by Franklin, Pernici, and Yuan (2005) during the period of 1997 to 2001. The retention rate in Louisiana, which had been holding at 8 percent, rose to 11 percent with the advent of high stakes testing. In the last five years it has held steadily at around 10 percent and is currently at 9.6 percent. Also, their research revealed that African-American students were held back at a higher rate than other ethnic groups. Finally, in contrast,

[i]n looking at the impact of high stakes tests on student achievement in gatekeeper grades, Roderick et al. (2002) examined school records of students in Chicago. The researchers found students with the lowest academic skills showing the largest gains on reading test performance in grades 3 and 6. In grade 8 all groups of students showed positive testing effects for reading; however, the opposite was true with mathematics. Higher achieving students showed the greatest gains in mathematics test performance in grades 6 and 8. For grade 3, high-risk students showed the greatest gains under the high stakes testing policy (Mancuso, 2004, p. 50).

Hamilton & Stecher (2004) summarize the potential effects of high stakes testing in

Table 2.3.

#### Potential Effects of High Stakes Testing

Many of the negative effects that were noted in Table 2.2 came to light in a book written by Dale and Bonnie Johnson (2002). This book, *High Stakes: Children, Testing, and Failure in American Study*, was the result of a year long study that took place in a rural section of Louisiana. The Johnsons left their professor positions at the University of Louisiana at Monroe

Table 2.3:

Potential Effects of High Stakes Testing (Hamilton & Stecher, 2004, p. 86).

Types of Effects	Effects on Students	Effects on Teachers	Effects on Administrators
Positive Effects	<p>Provide students with better information about their own knowledge and skills</p> <p>Motivate students to work harder in school</p> <p>Send clearer signals to students about what to study</p> <p>Help students associate personal efforts with rewards</p>	<p>Support better diagnosis of individual student needs</p> <p>Help teachers identify areas of strength and weakness in their curriculum</p> <p>Help teachers identify content not mastered by students and redirect instruction</p> <p>Motivate teachers to work harder and smarter</p> <p>Lead teachers to align instruction with standards</p> <p>Encourage teachers to participate in professional development to improve instruction</p>	<p>Cause administrators to examine school policies related to curriculum and instruction</p> <p>Help administrators judge the quality of their programs</p> <p>Lead administrators to change school policies to improve curriculum or instruction</p> <p>Help administrators make better resource allocation decisions, e.g., provide professional development</p>
Negative Effects	<p>Frustrating students and discourage them from trying</p> <p>Making student more competitive</p> <p>Cause students to devalue grades and school assessment</p>	<p>Encourage teachers to focus more on specific test content that on curriculum standards</p> <p>Lead teachers to engage in inappropriate test preparation</p> <p>Devalue teachers' sense of professional worth</p> <p>Entice teachers to cheat when preparing or administering tests</p>	<p>Lead administrators to enact policies to increase test scores but not necessarily increase learning</p> <p>Cause administrators to reallocate resources to tested subjects at the expense of other subjects</p> <p>Lead administrators to waste resources on test preparation</p> <p>Distract administrators from other schools needs and problems</p>

for one year and volunteered to teach at an elementary school. In that time the authors chronicled teachers' experiences through observations and their own experiences.

While reading this book, the reader is left with three impressions that are apparent concerning the impact of high stakes testing on teachers. These three impacts were that teachers' classroom practices were affected; teacher morale/commitment was altered, and pressure on teachers from outside sources was constant.

#### Classroom Practices, Including Test Preparation

“Q: With such an emphasis on testing, won't teachers 'teach to the test' and ignore other topics important to the overall education of students?

A: It is a fact that tests drive instruction. However, if the tests measure what students should know and be able to do, then it is appropriate that teachers incorporate LEAP21-type work into their daily teaching.”

The above is a quotation from “Reaching for Results: LEAP 21 (LDE, 2001, p. 21).

Despite the acceptance of the State regarding the appropriateness of teaching to the test, many critics of high stakes testing have found “teaching to the test” to be detrimental to the student.

“Teaching to the test” can be defined as a concentration on skills and activities that increase test scores with little concern for the amount of knowledge attained (Langenfeld, Thurlow, & Scott, 1997).

Teaching to the test has altered the priorities of teachers. Teachers felt they not only lost control over the way they teach, but that the way they were being required to teach contradicts what they learned as professionals. They felt the intensification of a test-oriented curriculum was occurring at the expense of the academic, psychological, social, and emotional needs of students (Mathison & Freeman, 2006, p. 56).

Other research came to the same conclusion that the standards compelled through high stakes testing narrows curriculum to basic skills and test driven content (Smith, 1991; Haney, 2000). Lazear (2006) reported that high stakes testing induced teachers to focus only on the test

and ignore untested aspects of knowledge. Teachers perceive that they have little control over setting the curriculum (Schultz & Oyler, 2006). Cimbricz (2002) found that teachers perceive that they have become testing coaches that are less likely to try innovative classroom practices. The author's interviews with teachers found that the perceived loss of control resulted in perceptions of anger and fear. McMillan, Myran, & Workman (1999) surveyed 722 Virginia teachers of which 80% reported that the State Standards of Learning (SOL) test had influenced their instruction forcing teachers to place greater emphasis on covering the content of the state test. Furthermore, Jacob (2005) found in studying large urban districts in the mid-west that science and social studies were less emphasized since they were not on the high stakes test. Dorn et al. (2005) found that geography was less emphasized. Pringle and Martin (2005) found that science was less emphasized. It was also found that the reading curriculum was confined and constricted by high stakes testing (Afflerbach, 2005). Canoy and Tut (2005) found that these methods led to rote memorization of facts.

On the other hand, Firestone, Mayrowetz, & Fairman (1998) found that although the tests may influence what teachers teach, they do not influence how they teach. Also 87% of Kentucky teachers surveyed believed that their state test, Kentucky Instructional Results Information Systems (KIRIS), had caused them to de-emphasize or ignore untested subject areas (Koretz, Barron, Mitchell, & Stecher, 1996). Stecher et al. (2000) also found similar results in Washington and Jones et al. (1999) in North Carolina. Herman & Golan (n.d.) reported that not only did teachers teach to the test but also they altered the sequence of their curriculum to ensure that content that was likely to appear on the test was covered first.

Stuart Yeh (2005) found that curriculum was narrowed in many schools by the excluding of subject matter not tested. Topics were ignored because they were unlikely to appear on the test. Learning was reduced to rote memorization of certain facts instead of higher order thinking

skills. Finally, an inordinate amount of time was spent on test preparation instead of actual learning.

Regarding too much time spent on test preparation, (Jones, Jones, & Hargrove, 2003) provided a study that states that 28% of teachers spend 60% of class time on test prep materials and 80% of teachers spend at least 20% of class time on test preparation.

These authors, also, found mixed reports on the use of projects, textbooks, lecture, and worksheets which increased or decreased due to high stakes testing.

Jones, et al. (2003) suggests the reason for conflicting results in the research is that much of it is based on interviews and self-reports. They present four factors to explain why research findings are mixed. Factors influencing the type and amount of change teachers make in their methodologies include the type of high stakes assessment, the type of professional development offerings, the subject area tested, and the level of achievement of the school (Mancuso, 2004, p. 44).

The Johnsons (2002) concluded that subjects such as science, social studies, art, drama, and music are abandoned and neglected. They advocate that teachers therefore need the freedom to teach and make decisions. “One teacher whispered, ‘Where will we find time to teach? All we do is test’” (Johnson & Johnson, 2002, p. 38).

An Arizona study stated that teachers placed less emphasis on non-tested subjects and shifted their time and attention to those subjects that were tested and addressed in high stakes testing (Smith et al., 1998).

The most important challenge the U. S. education system faces is not preparing students to do well on high stakes tests, but rather fostering 21<sup>st</sup> Century skills and knowledge in learners so that they are prepared to participate in our global, knowledge based civilization [and]... Current professional development that focuses on how to optimize teachers’ knowledge and skills within the current high stakes testing environment is tactically useful but strategically inadequate (Dede, 2004, pp. 12, 16).

According to some researchers high stakes testing has moved classroom instruction away from such constructivist activities as project based learning, open-ended learning environments, and solving real world problems. “The high stakes testing movement which has been bolstered

by the requirements of the No Child Left Behind Act of 2001, has inspired an environment in many schools that reinforces teacher-centered activities that use ‘lots of skills and drill’ activities” (Barksdall-Ladd & Thomas, 2000, p. 389). This atmosphere is contrary to the aims of student centered learning. “Nieto (2003) contends that high stakes testing may in fact be restricting pedagogy. For example, instructional practices indicative of student-centered learning, such as collaborative writing, science experiments, and thematic integrated units, are being suspended in some schools” (Grant & Hill, 2006, p. 20).

Teachers have not had enough power over the choice of curriculum ...however, the teachers in this study relate that previous to the state tests they had a lot more freedom, flexibility, and autonomy in deciding how they covered, delivered, and assessed the curriculum... For these teachers, the absence of choice and flexibility to make decisions about one’s teaching means that teaching as a profession is being threatened. Previous to the adoption of state mandated testing, teachers felt they played a vital role in making decisions about what to teach and how to teach to meet state or district standards (Mathison & Freeman, 2006, p. 51).

Also, in this study that took place in Upstate New York, the authors discovered that poor test scores did not only affect teachers but held strong consequences for schools and districts, as well. Therefore, the districts in the study all responded by adopting textbooks aligned to the format of the tests and created uniformity of teaching across grade levels. Other research that are related to high stakes testing’s effects on instructional strategies are listed and summarized by Langenfeld, Thurlow, and Scott (1997) in Appendix B.

#### Pressure

“Schools are now test-prep centers and woe be to those who don’t do enough prepping. Nikki, a beginner teacher, says, ‘I don’t know how much longer I can last. I just can’t get anything accomplished, and it just keeps piling on’” (Johnson & Johnson, 2002, p. 42). The book goes on to say that the teachers of this school perceive pressure from all sides. State department officials, officials from central office, the school site administrators, the local newspaper, and parents have combined to make their jobs much more stressful.

McNeil and Valenzuela (2000) found from their study that teachers report since the implementation of the Texas Assessment of Academic Skills (TAAS), a test similar to the LEAP, they have been ordered or pressured into altering teaching practices to include more test preparations. One school mandated that all classes include 20 minutes of test preparation.

In California, Alamillo and Viramontes (2000) reported that the top-down reform where teachers were limited in the pedagogical decisions that could be made in their classrooms led to a very tense environment throughout the school year.

Finally, Shepard and Dougherty (1991) found that teachers in their survey felt a great deal of pressure to improve test scores. In their study two districts that employed high stakes testing were surveyed, one from the southeast and the other from the southwest. In that study 79% of teachers surveyed reported that they felt “substantial” or “great” pressure from the district administration to raise test scores. Also, 66% of teachers reported that the media and newspaper were a source of great pressure.

Results of a survey in Maryland included 88% of teachers reporting that they believed that they were under “undue pressure” to improve teacher performance. When a similar question was put to Kentucky teachers, 98% reported that they also were under “undue pressure” (Koretz, 1996).

#### Teacher Morale and Commitment to the Profession

Evans (1998) defined morale as a state of mind decided by the person’s expectation of how far their needs are satisfied and how much those needs meaningfully affect that person’s total job situation. Halsey (1995) reported that educators had low morale based on poor working conditions and low status.

Young (2000) conducted a study in Hawaii that reported that teacher morale was related to job satisfaction. His findings provided evidence that “when teachers are distressed and the



educational system is perceived as letting them down, consequences such as teacher burnout, absenteeism, and attrition can result” (Kiziltepe, 2006, p. 146).

Maslach, Schaufeli, and Leiter (2001), in an extensive study of burnout, found the relationship of inefficacy to the other aspects of burnout is somewhat more complex. In some instances, it appears to be a function of either exhaustion, cynicism, or a combination of the two...The literature [Guglielmi & Taltrow, 1998] generally supports the hypothesis that occupational stress and teacher burnout are associated with poor health in teachers affecting the learning environment and interfering with the achievement of educational goals because they lead to teacher’s detachment, alienation, cynicism, apathy, absenteeism and ultimately the decision to leave the field (Kiziltepe, 2006, p. 147).

Due to the pressure and lack of autonomy in making classroom decisions, Johnson and Johnson (2002) reported a decrease in the morale of the teachers. As a result of the pressures related in the book, seventeen teachers sought jobs elsewhere.

In June 2001, Pam Porter announced her decision to leave Redbud to take a principalship in Arkansas. Her right-hand faculty member, veteran kindergarten teacher Carolyn Kesslem, will leave Redbud to teach at a university lab school. Both Pam and Carolyn had been at Redbud Elementary School for more than twenty years. Several Redbud teachers have accepted teaching positions in different districts, and five more teachers are looking for other jobs (Johnson & Johnson, 2002, p. 191).

In New York, which employs the New York State Regents Examinations, Bishop and Mane (2001) found that morale was enhanced by testing for many teachers who sought a challenge. Although, they admit that some teachers did quit.

However, other studies indicate that high stakes testing increases stress and decreases morale among teachers (Barksdale-Ladd & Thomas, 2000). “According to Jones et al. (1999), more than 77% of teachers surveyed indicated decreased morale; in addition 76% reported that teaching was more stressful since the implementation of the North Carolina state – testing program” (Pedulla et al., 2003, p.24).

Experts consider stress a serious matter and something that is needed to be kept to a minimum because of its negative influence on teacher motivation (Dörnyei, 2001). However, Young (2002) noted that despite the importance of a motivated teaching staff, these individuals

that dictate policy usually neglect or ignore teachers' needs. Policymakers often fail to consult teachers about new policies that may result in teacher morale problems. Confirming this belief, the Organization for Economic Cooperation and Development (1989) "reported that an uncommitted and poorly motivated teaching body will have disastrous effects for even the best of intentions for change" (Kiziltepe, 2006, p. 148).

Furthermore, Hoffman, Assaf, and Paris (2001) suggested that Texas' accountability system led to excellent teachers quitting. Currently in Jefferson Parish, the focus of this study, there is a shortage of 43 teachers at the middle school level. The reasons have not been researched.

Webb (2006) reports that a high level of stress among teachers led to demoralization because they were working under the threat of dismissal. In the years from 1991-2001, more teachers nationally have left the profession than have entered, with only 20% leaving because of retirement (Ingersoll, 2001). "Attrition rates for beginning teachers are particularly steep, as nearly one third will leave teaching within their first 5 years (Ingersoll, 2001). While several factors contribute to teachers' heightened stress, a significant amount of stress results from high stake accountability systems" (Webb, 2006, p. 2). This same thought is echoed by Sirotnik (2002, quoted in Webb, 2006, p. 2) who reported that "evidence is emerging about teacher demoralization and attrition as a result of frustration with the overemphasis on mandated testing for high stakes accountability purposes (p. 2)".

Mathison and Freeman (2006) note that the presence of high stakes testing has changed the nature of the job, adding an enormous amount of stress. These authors quote Quick et al. (1997) in that "performance increases with increasing stress up to an optimum point, and then the stress load becomes too great, resulting in depressed performance" (p. 4).

Teacher autonomy in classroom practices, perceived pressures, and morale/commitment

will have a great impact on the future direction of the Jefferson Parish school system with lasting long term effects. This study will provide insight into the current perceptions of those closest to the action.

#### A Review of the Literature on Teachers' Perceptions of Testing Programs

There are many studies concerning high stakes testing and their effects on schools, teachers, and students. Some of these studies gathered information through administrators and teachers by utilizing surveys, interviews, and classroom observations. The current research on teachers' perceptions is focused on four areas.

The review of current research on teachers' perceptions of state testing programs is organized around four main topic areas: (a) impact on classroom practices in terms of the content of instruction and the strategies used to deliver instruction, (b) the pressure to prepare students for the state test, (c) impact on teacher and student motivation and morale, and (d) views of accountability (Abrams, Pedulla & Madaus, 2003, p. 2).

In the area of the impact on classroom practices, McMillan, Myran, and Workman (1999) surveyed 722 Virginia teachers. In Arizona, Smith, Edelsky, Draper, Rottenberg, & Cherland (1998) utilized classroom observations and interviews of teachers in two elementary schools. Both of these studies found similar results. They found that teachers focused their teaching time in the classroom on subject matter that was tested and ignored or minimized material that the test did not cover.

In contrast, a study in Kentucky (Koretz, Barron, Mitchell, & Stecher, 1996) found mixed results. The Kentucky research, that surveyed a sampling of 4<sup>th</sup> and 8<sup>th</sup> grade public school teachers across the state, also found that teachers did de-emphasize or ignore untested subject areas. In addition, however, they found that teachers increased their instructional emphasis on problem solving and writing.

Perhaps these differences in research findings are a function of the format the state test, since Virginia's tests are predominantly multiple choice while the state test in Kentucky at the time of the study was based on portfolios (Abrams, Pedulla, and Madaus, 2003, p. 19).

In Maryland, Koretz, Mitchell, Barron, & Keith (1996) conducted a mail-in survey of 186 5<sup>th</sup> and 8<sup>th</sup> grade teachers and then followed it up with interviews of 112 principals and 224 teachers. Both Maryland and the Kentucky research revealed that teachers overwhelmingly felt undue pressure due to the testing.

In North Carolina, Jones et al. (1999) surveyed 470 elementary teachers across the state of which 236 respondents returned the survey. They found, along with the Kentucky and Maryland studies that testing had adversely affected teacher and student morale.

Finally, one study conducted by the National Board on Educational Testing and Public Policy (Pedulla et al., 2003) utilized an 80-item survey that employed a four-part Likert scale. This research was unique in that it involved many states with low stakes testing. 12,000 teachers were mailed surveys. 4,195 or 35% returned the surveys that covered all four of the areas of focus mentioned earlier. This study discovered that regardless of low or high stakes testing, teachers believed that they were forced to change their classroom practices for the worse. Although, teachers under high stakes testing were more likely to report spending more time on testable material and test preparation; the teachers from states employing high stakes testing felt greater pressure and lower morale. The results led the authors to conclude that the states needed to “refocus education policies to place greater emphasis on supporting and improving teaching and learning, rather than relying on a system of rewards and sanctions to spur change in the classrooms” (Abrams, Pedulla, and Madaus, 2003, p. 27). Other research that are related to high stakes testing’s effects on attitudes and motivation are listed and summarized by Langenfeld, Thurlow, and Scott (1997) in Appendix B.

### Mixed Methods Research Design

Many of the studies that have been accomplished recently have focused on using either quantitative or qualitative methods. Given the complexity of this research problem other methods

should be utilized to produce a complete picture. A picture that will produce answers that inform more than the simple numbers in a quantitative sense or words in a qualitative sense. A solution to this problem would be to utilize a mixed methodology research design.

Tashakkori and Teddlie (2003) described mixed methodologies as the “third methodological movement” (p. ix). This methodology can be defined as follows:

Mixed methods research is a research design with philosophical assumptions as well as methods of inquiry. As a methodology, it involves philosophical assumptions that guide the direction of the collection and analysis of data and the mixture of qualitative and quantitative approaches in many phases in the research process. As a method, it focuses on collecting, analyzing, and mixing both quantitative and qualitative data in a single study or series of studies. Its central premise is that the use of quantitative and qualitative approaches in combination provides a better understanding of research problems than either approach alone (Creswell & Clark, 2007, p. 5).

Creswell and Clark (2007) describe the evolution of mixed methods research as taking place over four stages of development. The first stage, which the authors call the “Formative Period”, began in 1959 with Campbell and Fiske who advocated multiple types of quantitative data to study psychological traits.

The second stage labeled the “Paradigm Debate Period” took place in the 1970’s and 1980’s. During this period Rossman and Wilson (1985) called researchers who refused to mixed paradigms-purists. Those who adapted their methods to the situation were called situationalists. Finally, the authors labeled researchers as pragmatists who believed that multiple paradigms could be utilized to address research problems.

The worldview called pragmatism is closely associated with mixed methodologies. The focus [of pragmatism] is on the consequences of research, on the primary importance of the question asked rather than the methods, and multiple methods of data collection inform the problems under study. Thus, it is pluralistic and oriented toward ‘what works’ and practice (Creswell & Clark, 2007, p. 23).

The third stage of development was labeled the “Procedural Development Period”. During this period, attention shifted toward the methods and steps for designing a mixed methods research study.

During this period, Greene, Caracelli, & Graham (1989) authored the article, “Toward a Conceptual Framework for Mixed-Method Evaluation Designs”. This article was important to establishing the groundwork for mixed methods research design. This led to many authors such as: Morse (1991), Creswell (1994), Morgan (1998), and Tashakkori & Teddlie (1998) developing classification systems.

The fourth and current stage of development is labeled, “Advocacy as a Separate Design Period”. During this period, Tashakkori & Teddlie (2003) followed up their previous work by creating a comprehensive 768-page “Handbook of Mixed Methods in Social and Behavioral Research”. The book’s “26 chapters [were] devoted to controversies, methodological issues, applications in different discipline fields, and future directions” (Creswell & Clark, 2007, p. 16). Also, Johnson and Onwuegbuzie (2004) called for mixed methods to be considered as a separate legitimate design in educational research.

#### Recent Developments in Mixed Methods Research

In the last seven years the National Institutes of Health, the National Science Foundation, and the American Educational Research Association have provided guidelines or workshops concerning mixed methods research. Also, the first international conference specifically devoted to mixed methods research was held in July, 2005 at Cambridge University.

Many prestigious journals such as the *Annals of Family Medicine*, the *Journal of Counseling Psychology*, the *Journal of the American Medical Association*, and *Lancet* have devoted issues to mixed methods research. The first issue of a new journal named *The Journal of Mixed Methods Research* became available in January, 2007.

Mixed methodologies have evolved over the four stages presented to become a powerful research design. This combination of qualitative and quantitative methods has the potential to provide better understanding of the research problems. This better understanding is the result of

combining both quantitative and qualitative methods thereby utilizing the strengths of one method to offset the weaknesses in the other method. Mixed methods can answer questions that either qualitative or quantitative methods could not answer alone.

The classification of mixed methods design chosen for this study will be more fully described in Chapter 3. This mixed methods design will constitute three phases. The first phase will be devoted to instrument creation. The second phase will utilize the survey that was created in the first phase. This survey will produce quantitative data that will inform the third phase. In the third phase interviews will be conducted to create in-depth qualitative data.

### Why Use Mixed Methods?

Tashakkori and Teddlie (2003) advocate the pragmatist's view of doing what works when conducting research. Mixed methods research is very compatible to the pragmatist's view.

What is healthy about a pragmatic social science of mixed and multiple methods is the fact that this effort has opened up languages of social science. It allows a number of projects to be undertaken without the need to identify invariant prior knowledge, laws, or rules governing what is recognized as 'true' or 'valid'. Only results count! Nor do we require a single foundational discourse of 'research methodology' to warrant our activities. 'Rationality' need not be affixed to a single overarching method of inquiry, nor do we require that the belief in any method or mixture of methods requires 'justification' for the pragmatic interest to win out (Maxcy, 2003, p. 85).

In pragmatism there is a rejection of the concept of either-or choices of the quantitative and qualitative research methods and instead embrace both points of view.

Miles and Huberman (1994) also believe the quantitative-qualitative argument [was] unproductive and there are benefits to linking qualitative and quantitative methods including: confirmation or corroboration of each other; more elaborate analysis by providing richer detail; and initiating new lines of thinking by providing new insights (Lange, 2002, p. 74).

Russek and Weinberg (1993) found that by using both quantitative and qualitative data new insights were added. Their data collected regarding elementary school teachers were given a depth that neither type of analysis (qualitative and quantitative) could provide alone.

In this study mixed methodology will be utilized because the final conclusions will be much richer than results emanating from one approach or worldview. By using this method words and narrative (interviews) are used in this study to add meaning to the numbers collected (survey). Conversely, numbers added precision to the narrative.

This approach can answer a broader and more complete range of research questions because I am not confined to a single method. The strengths of one method (qualitative or quantitative) are used to overcome the weaknesses of the other method by using both in the study (Johnson & Onwuegbuzie, 2004).

Also, I believe that inferences made from mixed methods are more complete and accurate representations of what is actually occurring rather than a collection of inferences collected separately from either a qualitative or quantitative approach alone.

Mixed methods research provides...

...stronger evidence for a conclusion through convergence and corroboration of findings. Can add insights and understanding that might be missed when only a single method is used... Qualitative and quantitative research used together produce more complete knowledge necessary to inform theory and practice (Johnson & Onwuegbuzie, 2004, p.21)

Teachers are complex participants that may have important knowledge of events that cannot be answered through one method alone. A complete picture will need mixed methods.

Utilizing both inductive and deductive logic, the mixed methodology employed in this research will inform the research questions and possibly explore unforeseen important variables. The survey (quantitative) is exploratory to discover teachers' perceptions regarding high stakes testing in the areas of classroom practices, perceived pressures, and teacher morale/commitment.

The interview method (qualitative) will provide depth and breath to the survey. Also, it may alert the researcher to variables and concerns not addressed in prior research.



## Importance of the Two Independent Variables

The two independent variables that will be utilized for this study are socioeconomic status and schools' past success regarding the LEAP test. The socioeconomic status (SES) will be assigned to schools due to the percentage of students who are designated as "free or reduced lunch". The schools' past success regarding the LEAP Test will be ascertained from data collected from the Louisiana Department of Education.

Many studies have shown that the composition of the socioeconomic backgrounds of all the students attending a school can affect the students' achievement beyond the effects associated with the students' individual ability and/or social class... Willms (1992) has suggested several other advantages that schools with higher socioeconomic student bodies may have over those with lower socioeconomic students: (a) they have greater support from parents; (b) they have fewer discipline problems; (c) they have atmospheres that are more conducive to learning; and (d) they are more likely to attract and retain excellent teachers (Freeman, 1997, p. 56).

Socioeconomic status can have an impact on those students in a high stakes testing environment. Teddlie and Reynolds (2000) stated, "Drawing on Teddlie's review on 'context in school effects research' (Teddlie, 1994), it becomes clear that contextual effectiveness is indeed a viable part of educational effectiveness research. The most frequently studied context variables are: average socioeconomic status of students ..." (p. 295). Reardon (1996) discovered that high stakes tests resulted in a rise in dropout rates only in schools with a lower socioeconomic status. Also, Tuerk (2005) found that schools with a high socioeconomic status have better access to more highly qualified teachers. Waber et al. (2006) found that students from socially disadvantaged backgrounds perform poorer on high stakes tests than their more advantaged peers.

Herman and Golan (undated) found that:

Correlations show that socioeconomic status is significantly and negatively related to the following: school attention to test scores, teachers' attention to testing and planning their instruction, and overall time devoted to test preparation activities . . . testing is more influential and exerts stronger effects on teaching in schools serving more disadvantaged students (p. 57-58) Not all studies found a relationship between SES and performance on

high stakes tests. Corbett and Wilson (1991) found that SES played a “surprisingly weak” role in explaining differences between districts (Langenfeld, Thurlow, & Scott, 1997, p. 17).

It is hard to determine from the available evidence whether low or high SES schools have different effects on teachers’ perceptions in a high stakes testing environment. The results do suggest, however, that we need to take a closer look at this question to make sure that we are not using high stakes testing as a means widening gaps that already exist in our current educational system.

The other independent variable that will be explored concerns School Performance Scores (SPS). I could find no studies that compared teachers’ perceptions in schools with high SPS scores to schools with low SPS scores.

Does past success or failure regarding LEAP21 Testing affect teachers’ perceptions toward the test? By identifying similar schools with different SPS scores and administering the same survey followed by interviews, I plan to answer this question.

Some studies found that prior success was important to establishing positive attitudes toward testing. Spickerman (1970) found that a positive attitude toward mathematics was the direct result of prior success in math courses. Furthermore, Adami- Bunyard, Gummow, and Millazo (1998) found that student past success rate did have an impact on student motivation and attitudes.

Hair, Kraft, and Allen (2001) examined twelve of the top twenty schools with the highest SPS scores in Louisiana based on the 1999 School Performance Scores. These schools also had between 80% - 100% of their students on free/ reduced lunch. The authors found that these shared characteristics that may not have been present in the lower performing schools.

Characteristics such as:

1. A variety of approaches to professional learning are present

2. The principal and entire staff have a strong sense of efficacy
3. Instructional leadership and faculty collaboration are key
4. The schools spent a great deal of time and attention on data analysis
5. The entire faculty demonstrated great flexibility in trying different approaches to meet student needs
6. Standards-based instruction was pervasive throughout all the classrooms
7. All schools had excellent school-wide discipline
8. Student learning was the school's greatest priority.

If schools with high and low SPS scores have different characteristics, do those different environments affect teachers' perceptions regarding high stakes testing? Also, when confronted with success or failure regarding a school's SPS scores teachers may rationalize the results. This rationalization could be sweet lemon or sour grape rationalizations. Interviews will help to identify if this present.

According to the "rationalization postulate", people cope with events by bringing their judgments of desirability into congruence with judgments of likelihood, so that 'one's wishes about an event tend to be adjusted to one's expectations about it' (McGuire & McGuire, 1991, p. 7). The authors specify a symmetrical relation, such that: This adjustment includes both (a) a 'sweet lemon' rationalization such that an increase in [likelihood] should raise [desirability] and so raise the number of desirable consequences that the core event is perceived as promoting and the number of undesirable consequences it is perceived as preventing; and also (b) a 'sour grapes' rationalization such that a decrease in [likelihood] should decrease [desirability] and so raise the number of undesirable consequences that the core event is perceived as promoting and the number of desirable consequences it is perceived as preventing (McGuire & McGuire, 1991, p. 7) ...The rationalization postulate holds that people will even embrace and adapt to unwanted outcomes by enhancing the subjective value of an event as it becomes more likely to occur. This is the most counter-intuitive aspect of the rationalization postulate: that people will even justify outcomes that are contrary to their own wishes and interests (Kay, Jimenez, & Jost, 2002, p. 303).

## Summary of Chapter 2

A detailed review was provided by this chapter that included a thorough review of the literature regarding high stakes testing and its effects on the educational process. It also explored

high stakes testing's effects on teachers' classroom practices, perceived pressures, and morale/commitment.

The review of the literature began with an explanation of the term, "high stakes testing" and its possible consequences. To fully understand the concept a history of high stakes testing both nationally and locally was introduced. This history included testing's path through the state legislatures and the courts.

Given the controversial nature of the subject matter, arguments for and against high stakes testing were presented to further illuminate the focus of this study. Also, effects of high stakes testing on classroom practices, both positive and negative, were explored.

The independent variables pertinent to the study, socioeconomic status and SPS scores, were fully presented. The study's dependent variables, classroom practices, perceived pressures, and morale/commitment, were also explored. Then, a review of the literature regarding teachers' perceptions of high stakes testing was examined.

This study uses a mixed methods design. Therefore, this research method was presented and explained.

The literature presented in this chapter was designed to illuminate the rationale for studying such an important, far-reaching concept as high stakes testing and its effects on teachers' perceptions regarding the educational process. In recent years its use has been escalated because of the No Child Left Behind Act; its impact needs to be fully understood. It is hoped that the results of this present study will make a contribution toward expanding the literature regarding teachers' perceptions regarding high stakes testing.

## CHAPTER 3: METHODOLOGY

### Introduction

As stated in Chapter 1, the purpose of this study is to assess the perceptions of Jefferson Parish teachers toward LEAP21 Testing and how high stakes testing affects school improvement. Currently, schools are evaluated based on their school performance scores. School Performance Scores (SPS) are based upon students' test scores on LEAP and iLEAP Tests, as well as their attendance records. Comparisons will be made between elementary schools with a high percentage of students that utilize free and reduced lunch (Lower SES) and those schools that have a lower percentage (Higher SES). Comparisons will also be made between elementary schools with low SPS scores and those elementary schools that have a history of achieving high SPS scores.

The study of the effects of high stakes testing has received a great deal of attention since the inception of No Child Left Behind. However, since this focus of study has only recently emerged, this research is limited. This study will utilize additional quantitative data to identify the actual effects of high stakes testing in the areas of morale/commitment, pressure, and classroom practices. This study will also explore if past success regarding SPS scoring affects the educational communities' perceptions toward high stakes testing. Also, the qualitative data will give further understanding and depth to the quantitative data.

This study utilizes both quantitative and qualitative research methods that are divided into three phases. The phases have the following goals: (a) to create an instrument (survey) that will explore teachers' perceptions regarding LEAP21 Testing; (b) to administer a survey to teachers in a purposive sample of schools with high and low SPS scores and schools with different levels of students that utilize free and reduced lunch (low and high SES), for the purpose of identifying differences between these categories of schools regarding a set of three dependent variables; and

(c) to identify teachers from the data collected from the surveys, and to initiate interviews with them to more fully understand the effects of high stakes testing on school improvement.

The survey method will be utilized to collect knowledge concerning the three key variables that are: classroom practices, perceived pressures, and teacher morale/commitment. The goal will be to have all teachers at each school complete the survey. The surveys will be left with a lead teacher and collected by the same lead teacher and put in a manila folder to be collected by the researcher.

Although Phase III (interviews) will be distinct from Phase II (surveys), it will be informed by the answers to the survey. The questions for the interview will be created after gathering the results of the survey. Interviewees will be chosen from the same schools that will be surveyed. The interviews will be voluntary.

In this chapter, I will first reintroduce the questions that guided the study and the design that this research utilized. Next, the phases of this study will be explored.

In Phase I, important terms will be defined. Also, the process of creating the survey will be presented.

In Phase II, the independent variables will be reintroduced. Also, in this phase sampling procedures and administration of the survey are discussed.

In Phase III, the interview method will be explored. This information will be followed by the data collection procedures, methods of analysis, and the inference techniques used in this study. Finally, the role of the researcher in this study will be discussed.

### Research Hypotheses

The progressive nature of this study was necessary because of the little amount of literature concerning the LEAP21 testing. Therefore, the first and second phase of this research

led to the development of concrete, quantitative data. This data will be further explored and analyzed in the third phase.

In order to expedite the design of the study, a series of research hypotheses were followed. The research hypotheses that were used as a guide throughout the phases are listed below.

1. Jefferson Parish teachers from schools that produce high SPS scores will perceive that LEAP 21 testing has affected classroom practice more than teachers from schools that have scored poorly.
2. Jefferson Parish teachers from schools which produce low SPS scores will perceive that LEAP 21 testing has created pressure to spend more time on test preparation (teaching to the test) than teachers from schools that have achieved high SPS scores.
3. Teachers from schools with high SPS scores will indicate that they have a higher degree of commitment to the education profession than teachers from schools with low SPS scores.

### Research Questions

The preceding research statements are confirmatory statements that will expand on knowledge extracted by a survey. The questions that I present now will be informed by answers that are retrieved by this survey and throughout the study and addressed in Phase III. The answers to the following questions will be generated by qualitative interviews.

1. How does test preparation (teaching the test) affect teachers' instructional planning, learning strategies, and curriculum content and to what extent?
2. How much time do teachers perceive that students spend on test preparation and how does that amount of time compare to the time spent on instruction?

3. What effect does testing have on an educators' sense of professionalism and pride in their work? How does high stakes testing affect motivation in general?

### Design For The Study

To accomplish the stated goals of the study a mixed methodology will be utilized. "A mixed methods study involves the collection or analysis of both quantitative and/or qualitative data in a single study in which the data are collected concurrently or sequentially, are given a priority, and involve the integration of the data at one of more stages in the process of research" (Creswell, Clark, Gutmann, & Hanson, 2003, p. 212).

In this case, both quantitative methods (surveys) and qualitative methods (interviews) will be utilized. These methods will be sequential with both phases of the study (Phase II – surveys) and (Phase III – interviews) being distinct. However, Phase II will inform Phase III.

Steven Miller (2003) believed the QUAN→QUAL research design superior to other designs. He saw great strength in its flexibility that a researcher can go in three directions to best answer the research questions.

[F]irst, the qualitative analysis is a way of supporting the original quantitative relationship, while the second suggests that there is another dimension beyond the quantitative that ought to be explored. A third dimension may be called 'emergent' and describes a situation in which a given statistical analysis is not statistically significant, but one then pursues the qualitative in some sense to 'explain' why there was no statistical significance (Miller, 2003, p. 442).

The mixed methods design that I will utilize will have three phases (instrument creation, surveys, and interviews). To explain the last two major phases (II and III) I will utilize a Sequential Mixed Model design. "The distinguishing attribute of the sequential mixed design is that the second phase of the study emerges as a result of, or in response to, the findings in the first phase" (Tashakkori & Teddlie, 2003, p. 687).



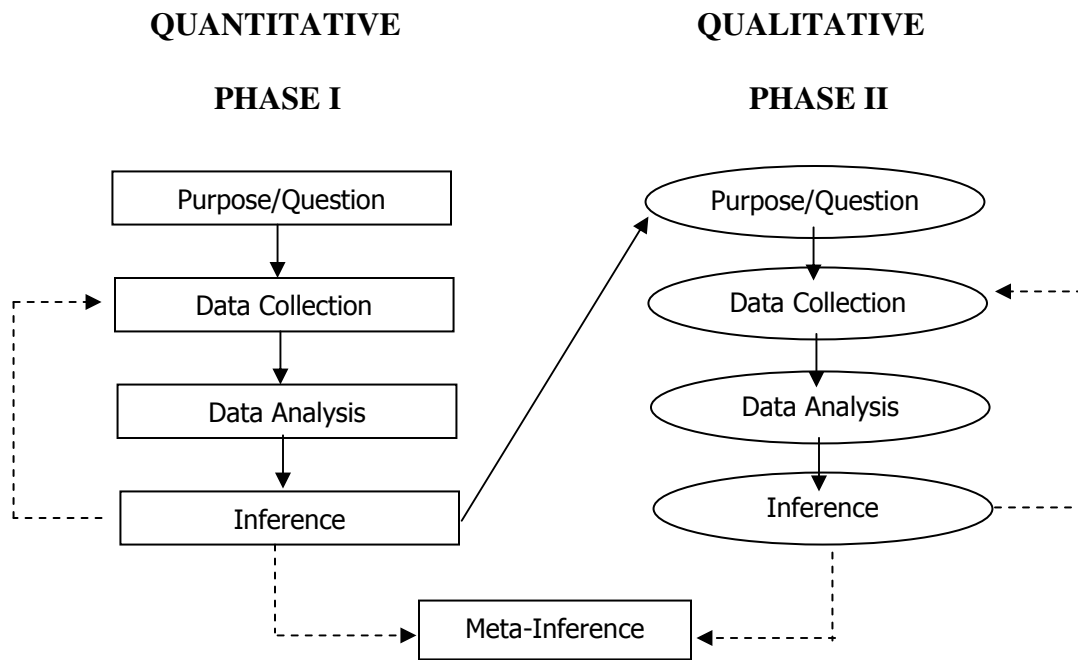


Figure 3.1:

QUAN – QUAL Methodology for Phases II & III

Phase I Methodology for Study: Instrument Development

Phase I of this study involved instrument generation and validation. The survey for this study was created by examining the previous research discussed in the Literature review section and guided by the stated research questions. From this information a survey instrument was created. To ensure the trustworthiness of this survey instrument a peer review was utilized and a pilot study using the instrument was performed.

The purpose of a survey is to generalize from a sample to a population so that inferences can be produced about a characteristic or attitude of the larger population. The advantages to using this quantitative method are its inexpensive design due to identifying attributes of a larger population from a small group of people and rapid turnaround in data collection (Babbie, 1990; Fowler, 1988).

A survey is an excellent way to gain knowledge from the target subjects of this study. It will allow the teachers to answer the questions comfortably and therefore more truthfully since their answers will be held anonymously (Patton, 1998).

There are disadvantages to the use of a survey which include: not getting an in-depth picture of why features are there or not or why stakeholders hold different perspectives; sometimes respondents will want to portray themselves in a better light; and the problem of bias in that the people who respond to surveys are usually the extremes of the continuum, most opinionated, generally better educated, wealthier, and less representative of minority groups” (Lange, 2002, p. 78).

Steps will be taken to address these concerns. Also, the surveys will be followed by interviews that support or contradict the information gathered during the surveys.

This study will utilize a modified instrument that is created by assembling the components of several instruments. I utilized a review of the literature to create this survey. Once the areas of classroom practices, perceived pressures, and teacher morale/commitment were identified as key features to teachers’ concerns then items that were related from other surveys were identified and grouped under these three areas of concern.

Besides the review of the literature the survey was greatly influenced by Lorrie Shepard and Katherine Dougherty’s (1991) questionnaire that asked teachers’ perceptions concerning high stakes testing. This questionnaire was administered in two large districts employing high stakes testing, one in the southeast and one in the southwest. This reputable study utilized a “[f]actor analysis to check on the validity or meaningfulness of the survey items” (Shepard & Dougharty, 1991, p. 4).

Also, this study’s survey was influenced by the survey research of Pedulla, et al. (2003). These authors conducted a national survey of teachers. This study surveyed 4,195 teachers across the United States for the National Board of Educational Testing and Public Policy.

The survey was based, in part, on other surveys used in Arizona (Smith, Nobel, Heinecke et al., 1997), Maryland (Koretz, Mitchell, Barron, & Keith, 1996), Michigan (Urda & Paris, 1994) and Texas (Haney, 2000), as well as on the National Science Foundation

(NSF) study of the Influence of Testing on Teaching Math and Science in Grades 4-12 (Madaus, West, Harmon, Lomax, & Viator, 1992) and a study of the Effects of Standardized Testing (Kellaghan, Madaus, & Airastan, 1980) (p. 16).

The survey employed a continuous Likert scale of five responses. These responses ranged from (1) Strongly Disagree to (5) Strongly Agree. Shepard and Dougherty (1991) were the source of questions 9, 12, 13, 14, 18, 19, 20, and 21. These questions covered the areas of perceived pressure concerning teachers, teachers' morale/commitment, classroom practices, and school improvement. Pedulla et al. (2003) influenced questions 2, 10, 11, 15, 16, and 17. These questions covered the areas of demographics, perceived pressure concerning teachers, teachers' morale/commitment, and classroom practices. I, Charles (2007), am responsible for creating the demographic questions 1, 3, 4, and 5. I also created questions 6, 7, 8 concerning teachers' morale/commitment but was heavily influenced by the research of Johnson and Johnson (2006). The source of the 22 questions included in this study is arranged by subject area and source in the table below.

Table 3.1:

Sources of the Questions Included in the Survey Used for this Study

Source	Demographics	Pressure	Morale/Commitment	Classroom Practice	School Improvement
Shepard & Dougherty (1991)		18, 19	9	12, 13, 14	20, 21
Pedulla et al. (2003)	2	16, 17	10	11, 15	
Charles (2008)	1, 3, 4, 5		6, 7, 8		22

The content validity of the instrument used in the current study was enhanced through the use of a peer review. I received input from four individuals that currently hold administration level positions in the Jefferson Parish School System and have attained a Ph. D. These experts

helped provide feedback concerning the content and clarity of the survey items. This feedback was used to improve these survey items.

### Pilot Study

Upon completion of the survey development, the survey was piloted to a group of teachers from two middle schools with similar free and reduced lunch percentages (SES) but different histories of success concerning the LEAP21 Testing. These teachers' feedback was incorporated into the final instrument. Piloting this survey was important to establish construct validity of this instrument and enhance the questions, format, and the Likert scale. This step was necessary because “[w]hen one modifies an instrument or combines instruments in a study, the original validity and reliability may not hold for the new instrument, and it becomes important to re-establish validity reliability during data analysis in a survey study” (Creswell, 2003, p.158).

A factor analysis is defined as a “technique based on how well various items are related to one another and form factors. Each factor represents several different variables, ... In using this technique, the goal is to represent those things that are related to one another by a more general name, such as a factor” (Salkind, 2004, p. 300).

Using Maximum likelihood analysis with Varimax rotation with Kaiser normalization, an exploratory factor analysis was conducted to ascertain simple patterns in the themes of relationships among the variables. Factor rotation was utilized “[s]ince the original loadings may not be interpretable, it is usual practice to rotate them until a ‘simpler structure’ is achieved. The rationale is very much akin to sharpening the focus of a microscope in order to see the detail more clearly” (Johnson & Wichern, 1992, p.419). Factor loadings on all the questions were guided by the work of Comrey (1973) who suggested the following measures: loadings > 0.71 excellent, loadings > 0.63 very good, loadings > 0.55 good, loadings > 0.45 fair and loadings > 0.32 poor.

This analysis was conducted using the data collected from the original surveys administered to two middle schools in Jefferson Parish. In the first school 30 teachers were surveyed. In the second school 36 teachers were surveyed. Since questions one through five were demographic questions they were eliminated from the factor analysis. The analysis addressed questions 6 through 21.

The first run of the data found six eigenvalues greater than one indicating six factors were present. Historically, eigenvalues less than one are considered trivial and are not considered for analysis (Hatcher, 1994). Observing the rotated factor matrix, the first three factors regarding the loading coefficients presented diverse results and no immediate interpretation of those factors evident. However, the factors four, five, and six did produce interesting results. Factor six produced only one very high loading coefficient (.955) for question six indicating that factor six is really representing question six. It was also discovered that factor five had only two questions (20 and 21) that had a rating of good or better. Therefore, factor five is representing questions 20 and 21. Factor four had produced a single very high loading coefficient for question 10 (.940). After consulting my peer debriefers, there was some concern whether question six and ten should be included in the final analysis. The factors four and six indicate that these two questions operate alone so they were dropped in the analysis. Questions twenty and twenty-one do not relate to the variables under study. These were removed as well and they appeared to operate as single coefficients.

After removing these four questions the factor analysis was run again in hopes to identify some more intuitive factors. The second run of the data found only three eigenvalues greater than one. Looking at the rotated factor matrix, factor one has high loading coefficients on questions 16 through 19f. The factor loadings for these questions ranged from .409 to .883. So factor one is corresponding to the pressure variable. Factors two and three were more diverse – there was a

high loading coefficient on questions seven and nine, however, we do not have a high loading coefficient on question eight otherwise that would have been the commit variable. Factor three had relatively high loading coefficients on questions eleven through fifteen, but additional high loading coefficients on 16 and 19f, however, progression was made toward isolating the commit variable and the classroom practice variable.

At this stage of the investigation factor one, the pressure variable did seem to be isolated. There was question about the content of the commit variable being negatively worded. I made the changes on the questionnaire to questions seven and eight. Then a second set of data was collected from a third pilot school. In that school 38 teachers were surveyed. Another factor analysis was run this time there were five eigenvalues that were larger than one indicating five factors present. In this new pilot data set I observed loading coefficients in the rotated factor matrix (Table 3.2), it was observed that the first factor had high loading coefficients on questions seven, eight, and nine (.705, .718, and .840) indicating that factor one represents the commit variable. In factor two, questions 11, 12, 13, 14, and 15 have high loading coefficients (.729, .702, .581, .721, and .609) indicating that factor two is representing the classroom practice variable. Factors three, four, and five were diverse and a poor representations of the pressure variable.

To analyze the results of this third pilot study the peer debriefers were consulted. It was decided that the first two factors (commit and classroom practice) had been isolated. It was also decided that the questions related to the variable pressure would remain the same for the following reasons:

- the pressure variable had been isolated in the previous pilot study
- these set of questions were taken from an established survey.

The survey instrument, after receiving feedback from the pilot group, can be viewed in Appendix C.

Table 3.2

Rotated Factor Matrix

	Factor				
	1	2	3	4	5
q7	.705	.918	.067	-.108	-.274
q8	.718	.279	.282	.157	.280
q9	.840	.096	.373	.145	.213
q11	.152	.729	.565	.251	-.053
q12	.156	.702	.555	.313	-.168
q13	-.030	.581	.045	.161	.048
q14	.272	.721	.165	-.005	.098
q15	.452	.609	-.182	.570	.012
q16	.687	.149	.240	-.008	.240
q17	.178	.212	.506	.013	.141
q18	.570	.309	.362	.266	.052
q19a	-.124	.040	.471	.867	.089
q19b	.141	.058	.047	.022	.723
q19c	.415	.067	.607	.188	-.041
q19d	.403	.190	.068	.481	-.393
q19e	.079	.331	.101	.691	.018
q19f	.365	.095	.556	.197	-.011

Note. It was observed that the first factor had high loading coefficients on questions seven, eight, and nine (.705, .718, and .840) indicating that factor one represents the commit variable. In factor two, questions 11, 12, 13, 14, and 15 have high loading coefficients (.729, .702, .581, .721, and .609) indicating that factor two is representing the classroom practice variable. Factors three, four, and five were diverse and a poor representations of the pressure variable.

Mixed Method Data Collection Procedures

I used the voice of educators as the starting point. I met individually with experts (educators) in the Jefferson Parish School System as a peer review. I used the peer review in conjunction with previous research to decide the content, sequence, and wording of the questions that make up the survey.

The survey is a combination of one open-ended and twenty-one close-ended questions. Open-ended questions are questions to which there is not one definite answer and teachers will be free to elaborate on their answers. The responses to open-ended questions can be very useful,

often yielding quotable material. These responses will also be very helpful in identifying teachers who will provide a rich, thick description of the phenomena being studied.

Closed-ended questions, on the other hand, have a finite set of answers from which the respondent chooses. The benefit of closed-ended questions is that they will be easy to standardize, and data gathered from closed-ended questions lend themselves to statistical analysis (Fink, 1995). I will disseminate the surveys in a manner chronicled in later sections. I will use the interviews to learn the issues in-depth.

The Interview Guide Approach will be utilized to develop the interview protocol in order to “increase the comprehensiveness of the data and make data collection somewhat systematic for each respondent. Logical gaps in data can be anticipated and closed. Interviews remain fairly conversational and situational” (Patton, 2002, p. 349). This approach will lead to the flexibility needed to fully inform the research questions.

Qualitative interview procedures can identify experiences in an attempt to understand or interpret phenomena in terms of the meanings that people bring to it (Denzin & Lincoln, 1994). Qualitative research assumes that reality is subjective and constructed by the participants involved in the study.

After I collect and conduct a preliminary analysis of the quantitative data, the stratified purposeful sample will be selected (explained fully in the Sampling section) and the interviews will be conducted. The advantages of the interview method that apply to this study are: providing depth to the data, allowing for probing, and improving the confirmation of the quantitative data (Johnson and Turner, 2003). A potential disadvantage that will be addressed in this study is a perceived lack of anonymity by respondents. Interviewees and respondents will be assured of their anonymity prior to beginning the interview.



Those teachers that will be interviewed will be contacted by phone at their respective schools. Those participants who responded positively to this request will then be scheduled for interviews. Interviews will be conducted at the teachers' schools by me and will last approximately 20 to 35 minutes. An Interview Guide of questions will be determined prior to the interview (see Appendix D). Some questions will be developed in accordance with the research questions; others will be determined after initial analysis of the quantitative data. Questions will be created to confirm the results of the quantitative phase of the study. Also, the interviews will allow for exploring the results of the quantitative data in more depth. The interviews will be recorded and fully transcribed.

During Phase III (interviews) I will interact with the participants to be interviewed. With the value-laden nature of this research, it is paramount that I identify my biases and values.

Maxwell (1996) believes qualitative research is concerned with... "understanding how a particular researcher's values influence the conduct and conclusions of the study. Validity in qualitative research is not the result of indifference, but of integrity" (p. 91). My experience with LEAP 21 testing as a teacher and administrator has not been a positive one. I plan to guard closely against bias.

The researcher will take detailed notes and also tape the interviews, which will be open to review to identify bias. Also, an interview protocol will be created and followed closely to ensure that each interview is similar. Finally, member checks will be utilized to ensure that the researcher accurately portrayed the perceptions of the teachers interviewed and not arbitrarily assumed his own experiences into their answers.

### Instrumentation Terms

The following instruments will be utilized in this study:

1. Peer Review - a process of subjecting the survey to the scrutiny of others who are

experts in the field. This process will help to create a more informed survey with more relevant questions.

2. Survey - a gathering of a sample of data or opinions considered to be representative of the opinions shared by the population of Jefferson Parish Public School teachers. The survey will be in a form containing a set of questions, which will address the hypotheses presented earlier.
3. Interview - a structured social interaction between a subject who is identified as a potential source of information by their answers to the survey and me, in which I will control the exchange to obtain quantifiable and comparable information relevant to the previously stated hypotheses.

#### Phase II: Quantitative Phase

The second phase will be a quantitatively driven project. This phase of the study will answer the three Research Hypotheses.

The independent variables throughout this study will be the SPS scores and socioeconomic status. The dependent variables will be the teachers' perceptions regarding classroom practices, pressure, and commitment. I believe that schools with low/high SPS scores or high/low socioeconomic status will affect teachers' perceptions of teaching practices, pressure and morale/commitment.

#### Determination of School Performance Score (SPS)

Prior successes' influence on perceptions was chronicled in Chapter 2. It was found that although there were not many studies available, the research that was presented found that prior success did have an impact on one's perceptions. It was also found that schools with high SPS scores may have different characteristics than schools with low SPS scores.

After the LEAP and iLEAP tests are administered to the children in the third week of March, each school is graded on an annual basis indicating the academic status of its students. This grade is called a School Performance Score (SPS). The SPS for each school is a composite index that utilizes indicators and weighting factors. The formula for calculating the SPS at the elementary school level is LEAP Test (60%), iLEAP Tests (30%), and Attendance (10%) (2006-2007 Louisiana State Education Progress Report).

Schools are assigned Performance Labels based on these SPS scores. A school must achieve a 60 or above in order not to be labeled “Academically Unacceptable” by the state. Using the Performance Labels as a guide, I initially assigned elementary schools that were labeled “One Star” or Academically Unacceptable” as schools with low SPS scores. Essentially, these are schools that have SPS scores at or below 79.9. Conversely, schools that are rated as “Two Stars” or above were classified as schools with high SPS scores. These are schools with SPS scores of 80.0 or above. This would split the elementary schools into two groups and was consistent with the fact that in 2004-2005 the average SPS score for Jefferson Parish for an elementary school was 78.18. However, I found that using these guidelines produced only one school that met the guidelines for the “High SPS/ Low SES” category. To rectify this I lowered the cutoff to the SPS score to 73 which allowed the inclusion of more schools into that category.

The data collected for this study came from the 2004-2005 school year. This was the last year of reliable information due to Hurricane Katrina. A report issued by RAND Education for the RAND Gulf States Policy Institute (Pane et al., 2006) found that the vast majority of students were displaced by the storm and most did not return to their original schools by the end of the year. Due to the hurricane more than a 25% of Louisiana’s 740,000 public school students were displaced. The study revealed that this was the largest displacement of students in the history of

the United States. In Louisiana 81% of the displaced students came from three parishes: Orleans, Jefferson, and Calcasieu.

Also, the study found that 45% of these students did eventually return to their original schools. This RAND study was done after many students enrolled somewhere else temporarily. Another 24% did not return to their original school but enrolled in another school within the state. The Department of Education chose to not release Jefferson Parish's scores for 2005-2006 on their website.

The SPS scores were found at the Louisiana Department of Education's website (Retrieved January 15, 2007 from <http://www.louisianaschools.net/lde/pair/ReportCards05pa/Principals/S026060.pdf>). Each school's SPS score was listed. Other demographic data was also retrieved for the survey from the "Principal's Report Card" which is also listed on the website. This information was later supported by the 2006-2007 unpublished data provided by a director in the Jefferson Parish School System. The 2006-2007 SPS scores were consistent with the criteria that were set forth previously. These data will be used to determine if teachers from schools that have a history of achieving high SPS scores have different perceptions regarding the LEAP21 test than teachers from low performing schools.

#### Determination of Socioeconomic Status (SES)

The research concerning the impact of a child's socioeconomic status regarding their academic performance is chronicled in Chapter 2. The result of the research was that a student's environment does have an effect on his/her performance in school (Coleman et al., 1966; Jencks et al., 1972; Freeman, 1997). Therefore, it is important to understanding that socioeconomic status be considered when comparing the perceptions of teachers from different schools.

Public schools in Jefferson do not maintain accurate data regarding the family income and the educational backgrounds of the parents who attend. Therefore, the method that will be utilized to designate a school's socioeconomic level is the data collected from the free/reduced lunch program (National School Lunch Program).

The National School Lunch Program [is] a federally assisted meal program that provides low-cost or free lunches to eligible students. It is sometimes referred to as the free/reduced lunch program. Free lunches are offered to those students whose family incomes are at or below 130 percent of the poverty level; reduced-price lunches are offered to those students whose family incomes are between 130 percent and 185 percent of the poverty level (retrieved from the United States Department of Agriculture website, on March 19, 2007).

The percentage of students involved in the free/reduced lunch program is maintained by the schools and reported to the Louisiana Department of Education and the Jefferson Parish Public School System's central office. Since a high percentage of students eligible for free/reduced lunch translate to a high number of parents near or below the poverty line, this data can be used to determine socioeconomic status (SES) of the school. In Louisiana, to calculate this percentage, the number of students that are enrolled in the free/reduced lunch program is divided by the total number of students attending the school (Crone et al., 1992; Freeman, 1997).

Records of a school's percentage of students that are enrolled in free/reduced lunch were retrieved from the Jefferson Parish Public School System's Title 1 office. This office is a liaison to the United States Department of Education that targets monies to the highest poverty schools by authorization of the Title 1 provision of the Elementary and Secondary Education Act of 1965.

In Jefferson Parish, until recently, only schools that had a percentage of students enrolled in free/reduced lunch at or above 75% were considered Title 1 (high poverty) schools. I will use this as a guide. The schools that are above or at 75% will be designated as low socioeconomic

level schools. Those schools which had 74.9% free/reduced lunch and below will be considered high socioeconomic level schools.

### Survey Instrument

The survey instrument (see Appendix C) utilized in this study was confirmatory in nature since it is being used to test the three hypotheses. As stated earlier, it is a culmination of previous surveys. Those surveys that contributed to this study's survey were Shepard and Dougherty (1991) and Pedulla et al. (2003). This survey consisted of 22 questions.

The first five questions are descriptive in nature and focused on teacher demographic information. The following are a list of items explored in this dependent variable group: school name, years of experience, grade level and subject or subjects taught, and how many classes taught in a day. These items are numbered 1 through 5 in the first section of the survey (see Appendix C).

Demographic information that was not included in the survey (but was retrieved from other sources that were discussed previously) will also be included in the context information. Examples of these data would be a school's socioeconomic status (free/reduced lunch percentage) and School Performance Scores.

Questions 6 through 10 address the effect of testing on teachers' morale/commitment, commitment, and job satisfaction. The answers to these questions utilized the Likert scale and ranged from (1) Strongly Disagree to (5) Strongly Agree. These questions explored the areas of tension, loss of control, job satisfaction, job abandonment, commitment, and unfairness and how these areas affect teacher morale/commitment.

Questions 11 through 15 focuses on classroom practices and how LEAP21 testing has influenced those practices. In this section changes in content and objectives are explored, along

with the school's focus on overall instruction. This section, like the previous sections, utilizes the Likert scale.

Questions 16 through 19 explore the perceptions of pressure felt by the teachers to increase scores on the LEAP21 test. These questions focus on where the pressure emanates and how it is focused. These questions use the Likert scale; however question 19 does not utilize the choices of "Strongly Disagree" to "Strongly Agree". Instead, question 19 lists six possible sources of pressure and asks the respondent to choose from (1) Almost No Pressure to (5) Great Pressure.

The last section of the survey (questions 20-22) explores teachers' perceptions toward whether LEAP21 testing has helped or hindered school improvement efforts. Questions 20 and 21 utilize the Likert scale with choices (1) Strongly Disagree to (5) Strongly Agree. The last question is an open-ended question that asks the reader to state their opinion concerning two positive and two negative consequences regarding the LEAP 21 testing. Responses to this question will be used to develop possible questions for the interviews in Phase III.

#### Mixed Methods Sampling Procedures

Kemper, Stringfield, & Teddlie (2003) state "[t]he sampling strategy should stem logically from the conceptual framework as well as from the research questions being addressed by the study" (p. 275). In order to accomplish this goal during Phases II and III, "Multilevel Mixed Methods" sampling will be utilized. The researcher chose this sampling technique because it "uses data collected from multiple levels within an organization or group so as to increase the power of inferences drawn from the data" (p. 286).

During Phase II, the researcher will work within this sampling framework, which involves the population of Jefferson Parish Public Schools. I will use Stratified Purposive sampling techniques to divide the purposefully selected target population into the following

strata: elementary schools that earned high SPS scores and elementary schools that had low SPS scores. I will select four schools from the former and four schools from the latter. These groups will be split once again into the following strata: elementary schools that have a high percentage of students in the free and reduced lunch program and those that have a smaller percentage. As explained before, researchers often use the free/reduced lunch program as an indicator of students from poor socioeconomic (SES) backgrounds.

This further division will create four cells (see Figure 3.2). Two schools will randomly be chosen from each of these four cells to create a sample of eight elementary schools. All teachers within these schools will receive the same surveys.

		Low SES	High SES
SPS Scores	Low	N=2 schools	N=2 schools
	High	N=2 schools	N=2 schools

Figure 3.2:

Sampling Procedures Surveys Low and High Socioeconomic Status (SES)

During Phase III, I will utilize Multistage Cluster sampling. This sampling technique allows the sample to be further reduced within the eight clusters (schools) to identify teachers who can add knowledge to the interview. Normally, the Multistage Cluster sample is random; however, it will add richness to the data if I identified teachers through their survey responses who can provide the most information for the questions under study. Four teachers will be interviewed from each cell for a total of sixteen teachers (Figure 3.3).



		Low SES	High SES
SPS Scores	Low	N=4 teachers	N=4 teachers
	High	N=4 teachers	N=4 teachers

Figure 3.3

Sampling Procedures Interviews Low and High Socioeconomic Status (SES)

Administration of the Survey

Once the sample is determined, the goal will be to have all teachers at each school complete the survey. Incentives, such as gift certificates will be utilized in order to help this objective. The surveys will be left with a lead teacher and collected by the same lead teacher and put in a manila folder to be collected by me. A cover letter explaining the purpose of the survey and directions will also be provided to the lead teacher.

Phase III: Qualitative Phase

With the information gathered from the second phase, I plan to initiate Phase III. In this third phase of this study I will utilize a qualitatively driven design. I plan to explore in-depth how school and classroom practices have been affected by high stakes testing. This phase of the study will answer the three Research Questions.

I plan to look at the effects of high stakes testing, LEAP21 testing in particular, from a teachers' perspective with the intention of gaining new insight. The rationale for incorporating a qualitative method and selecting a research design based on the paradigm of pragmatism is that it allows for the portrayal of the emic (insider's) point of view. This inquiry emanates from my intention to look at the issue of high stakes testing through the eyes of those being studied.

The ability of qualitative data to more fully describe a phenomenon is an important consideration not only from the researcher's perspective, but from the reader's perspective as well. "If you want people to understand better than they otherwise might, provide them information in the form in which they usually experience it" (Lincoln and Guba, 1985, p. 120). Qualitative research reports, typically rich with detail and insights into participants' experiences of the world, "may be epistemologically in harmony with the reader's experience" (Stake, 1978, p. 5) and thus more meaningful (Hoepfl, 1997, p. 49).

The qualitative method that will be utilized in Phase III is the interview method. The interview method that will be used is the interview guide approach.

An interview guide or "schedule" is a list of questions or general topics that the interviewer wants to explore during each interview. Although it is prepared to insure that basically the same information is obtained from each person, there are no predetermined responses, the interviewer is free to probe and explore within these predetermined inquiry areas. Interview guides ensure good use of limited interview time; they make interviewing multiple subjects more systematic and comprehensive; and they help to keep interactions focused. In keeping with the flexible nature of qualitative research designs, interview guides can be modified over time to focus attention on areas of particular importance, or to exclude questions the researcher has found to be unproductive for the goals of the research (Lofland and Lofland, 1984 quoted in Hoepfl, 1997, p. 52).

The overall design for the study is QUAN → QUAL. This design will lead to meta-inferences at the end of the study. These final meta-inferences are made on the basis of the confirmatory or disconfirmatory nature of the inferences in the two strands [phases] of the study (Creswell, 2002).

#### Mixed Methods Analysis

During the survey portion of the study (Phase II), I will utilize the SPSS computer software. The data retrieved from the Likert scales utilized in the survey will be imported into SPSS programs, which will generate means and standard deviations to represent the information. Finally, a One-Way and Two-Way Analysis of Variance with two levels (high SPS scores, low SPS scores) will be utilized to determine if the differences in the answers from teachers of schools with low and high SPS scores are significantly different, thereby testing Research Hypotheses #1-#3.

During Phase III (interviews), the purpose of the in-depth interviews is not to simply get answers to Research Questions #1-#3 but also to elicit accurate perceptions from those closest to high stakes testing. “At the root of in-depth interviewing is an interest in understanding the experience of other people and the meaning they make of that experience” (Seidman, 1991, p. 3).

Interviews will be taped and transcribed shortly thereafter. These transcripts will be read and reread, thereby creating codes and categories. These codes and categories will then be entered into programs utilizing the “Atlas.ti” software. The researcher will be constantly searching for patterns and themes.

The qualitative data analysis for this research will occur simultaneously with the collection of the data. It is ongoing process where the data read and reread always looking for patterns and themes. This method begins from the start of collection of the data. This data analysis can best be described as the constant comparative method (Glaser, 1978).

This method follows steps that fall into six areas:

1. Start collecting data
2. Search for key issues, recurrent events, or activities in the data that become categories of focus.
3. Collect data that provide several incidents of the focus categories while trying to see the diversity of the dimensions under the categories.
4. Write about the categories you are exploring, trying to describe and account for all the incidents you have in your data while constantly searching for new incidents.
5. Work with the data and emerging model to uncover basic social processes and relationships.
6. Employ coding and writing as the analysis focuses on the core categories” (Bogdan & Biklen, 1998, p. 67).

Coding, patterns, categories, and themes will be constantly employed to reduce the data into manageable entities. After the data has been coded and organized into categories, the next step will be to generate meaning across the interviews and how they address the research questions. The settings for the surveys and interviews will be at the teachers’ school sites to enhance comfort levels of the participants.

### The Mixed Method Inference Process

Concerns regarding content and construct validity were addressed as they pertain to the survey instrument. Content validity is defined as the belief that questions that are part of the instrument and data from these questions are representative of all possible questions that could be asked about the content being assessed (Creswell, 2002). Since the survey instrument utilized in this study was modified from two existing, published instruments, content validity is addressed. One of the surveys, Pedulla et al. (2003), reported that their survey questions' data yielded a Cronbach's alpha of .79. The other survey that contributed to this study's instrument, Shepard and Dougherty (1991), used a factor analysis to prove their survey's validity and meaningfulness.

Construct validity refers to whether the scores are consistent or measure what they intend to measure (Creswell & Clark, 2007). Construct validity will be addressed through the factor analysis performed on the survey. By examining the results of the pilot study of the survey I will be able to determine if the survey instrument measured the constructs it attempted to measure.

“Internal validity means that the investigator can only draw correct cause-and-effect inferences from the sample to the population if the threats are accounted in the design” (Creswell & Clark, 2007, p. 134). Campbell and Stanley (1963) identified eight threats to internal validity. These threats are: history, maturation, testing, instrumentation, statistical regression, differential selection of participants, mortality, and interaction effects. Many of these threats are addressed because of the nature of the instrument that I am using.

The fact that the instrument is disseminated to the teachers to take home and brought back a short time later will help mitigate threats such as history, maturation, mortality, and interaction effects. Furthermore, the instrument utilizes a Likert scale that is administered once. This should address the threats of testing, instrumentation, and statistical regression. Also,

differential selection of participants is addressed by surveying all the teachers at the selected elementary schools.

Also, to enhance the Internal Consistency Reliability of this study a Cronbach's Alpha was performed. "A Cronbach's Alpha is one of the most widely used diagnostic measure of the reliability coefficient that assesses the consistency of an entire scale of related questions. The measures range from 0 to 1. The generally agreed upon lower limit accepted for Cronbach's Alpha is .70 (Hair et al., 1998)" (Abington-Cooper, 2005, p. 97). The Cronbach's Alpha for this study was found to be .785.

I will enhance the Trustworthiness of the results that investigates the extent to which the researcher can persuade audiences that the findings are relevant. To accomplish this "rigor necessitates that researchers attempt to be fully accountable for their data collection, analysis, and interpretive methodologies. As noted by Onwuegbuzie (2000), such accountability implies that researchers continually strive to assess and document the legitimacy (e.g., credibility, trustworthiness, dependability, confirmability, transferability) of their findings" (Onwuegbuzie & Teddlie, 2003, p. 354).

Creswell (2003) states that the design utilized in this study enhance validation.

In a sequential approach, obtain themes and specific statements from participants in an initial qualitative data collection. In the next phase, use these statements as specific items and themes for scales to create a survey instrument that is grounded in the views of the participants. A third, final phase might be to validate the instrument with a large sample representative of a population (p. 221).

This method describes the pilot study that was conducted during the instrument development. This phenomenon can be found in Phase I.

After every interview, the transcription of the interview will be shared with the interviewee (Member Checks) to enhance credibility. Credibility shows us if the reconstructions of the researcher are "credible to the constructors of the original multiple realities" (Lincoln and

Guba, 1985). Credibility is based less on sample size than on the data obtained and on the analytical skills of the researcher (Patton, 1990). These member checks will be used to weigh the accuracy of the results by taking these specific descriptions back to the participants. These teachers then can voice whether they perceive that the notes taken were accurate. Credibility will be further enhanced by sharing all the steps of the process with my peers (Peer Debriefing) to ensure clarity. “This process involves locating a person (a peer debriefer) who reviews and asks questions about the qualitative study so that the account will resonate with people other than the researcher” (Creswell, 2003, p. 196).

Also, my transcripts will utilize Thick Description that will fully describe findings and interpretations to enhance Transferability. Transferability is the extent to which “the transferring of inferences from a specific sending context to a specific receiving context” (Tashakkori and Teddlie, 1998, p. 92). Transferability depends on the degree of similarity that exists between the situation in this study and a situation that it is transferred. “The researcher cannot specify the transferability of findings; he or she can only provide sufficient information that can then be used by the reader to determine whether the findings are applicable to the new situation (Lincoln and Guba, 1985)” (Hoepfl, 1997, p. 59). Through thick description of the findings I plan to transport the reader into the setting that will provide understanding.

In this chapter, the methodology of how the research questions will be addressed and how the study was designed was presented. The next chapter will present the analysis and interpretation of the data collected.

#### Researcher’s Role

During Phase II (surveys), the researcher’s role was minimal. Other than recruiting a lead teacher to hand out the surveys, I will not have contact with those teachers being surveyed.

During Phase III (interviews), the researcher will interact closely with the interviewees. Since, there is the possibility of exerting undue influence on those being interviewed; great pains will be taken to not interject personal opinion into the questioning. The steps to ensure that bias is not injected into the interviews were chronicled earlier in this chapter.

#### IRB and Jefferson Parish Public School System Approval

A written request for approval of projects that use human subjects was sent to the Institutional Review Board (IRB) at Louisiana State University for approval to perform this research. All the teachers and schools were sent approval letters prior their participation in the study (Appendixes E & F). IRB approval was received from the LSU Institutional Review Board, the issued IRB number for this study was #3715 (Appendix G). Jefferson Parish Public School System approval was also established (Appendix H).

#### Summary of Chapter 3

The methodology that was necessary to address the research hypotheses and questions was presented in Chapter 3. The hypotheses and questions that directed this mixed methods study utilizing three phases were included. Each phase was identified and the method in which the questions were to be addressed was explained. The phases had the following goals: (a) to create an instrument (survey) that will explore teachers' perceptions regarding LEAP21 Testing; (b) to administer a survey to teachers in a purposive sample of schools with high and low SPS scores and schools with different levels of students that utilize free and reduced lunch, for the purpose of identifying differences between these categories of schools regarding a set of three dependent variables; and (c) to identify teachers from the data collected from the surveys, and to initiate interviews with them to more fully understand the effects of high stakes testing on school improvement.

In Phase I, the survey instrument will be created. In Phase II, the surveys will be

disseminated and quantitatively analyzed. During the third phase, interviews will be conducted and qualitatively analyzed. All the steps and data analysis and collection techniques of all the phases were discussed in this chapter.

Descriptions of member checking, peer debriefing, and thick description were given to ensure the trustworthiness of the results of the study. The next chapter provides the outcomes of the quantitative part of this research.



## CHAPTER 4: QUANTITATIVE RESULTS

### Introduction

In this chapter, a brief description of the schools which participated in the study is presented. Due to the promise of confidentiality, the identities of the schools and participants of this study are not disclosed. The basic demographic information of the school sites in this study that were provided by the respondents and public documents are included in this chapter.

A sample of eight elementary schools was taken from the population of Jefferson Parish Public Schools. A stratified purposive sampling technique, described in detail in Chapter 3, divided the purposefully selected target population into the following strata: elementary schools that earned high SPS scores and elementary schools that had low SPS scores. These groups were split once again into the following strata: elementary schools that have a high percentage of students in the free and reduced lunch program and those that have a lower percentage. As explained before, researchers often use the free/reduced lunch program as an indicator of students from poor socioeconomic (SES) backgrounds. These actions created four groupings. Two elementary schools from the four groups were randomly selected for this study to create a sampling of eight schools. The criteria for creating the four groups are described in Chapter 3. The surveys were disseminated during the fall semester before LEAP 21 testing was conducted.

As indicated earlier, these surveys were created to produce quantitative data and inform the Research Hypotheses:

1. Jefferson Parish teachers from schools that produce high SPS scores will perceive that LEAP 21 testing has affected classroom practice more than teachers from schools that have scored poorly.

2. Jefferson Parish teachers from schools which produce low SPS scores will perceive that LEAP 21 testing has created pressure to spend more time on test preparation (teaching to the test) than teachers from schools that have achieved high SPS scores.
3. Teachers from schools with high SPS scores will indicate that they have a higher degree of commitment to the education profession than teachers from schools with low SPS scores.

Table 5.2 in the next chapter summarizes the quantitative and qualitative results of the data concerning the Research Hypotheses.

### Data Collection

Data collection for this research consisted of spending at least two days at each school site: one day handing out the surveys and one day interviewing teachers. Once the sample was determined, the surveys were disseminated to teachers at each school. Incentives, such as gift certificates were utilized in order to induce the completion of the surveys. The surveys were left with a lead teacher at each school site and collected by the same lead teacher, who then put the surveys in a manila folder to be mailed to me. A cover letter explaining the purpose of the survey and directions was also provided to the lead teacher (Appendix E). The composition of the surveys was covered in the previous chapter.

### Descriptions of Participating Schools

The Jefferson Parish Public School System's website boasts that it is one of the largest districts in the state and is nationally ranked in the top 100 for student enrollment. There are 84 schools located on the east and west banks of Jefferson Parish with a total enrollment of 52,367 students. Pre-kindergarten has 2,440 students with the remaining 49,927 in grades kindergarten through 12. Employees total 6,239, of which 3,494 are teachers. Jefferson Parish has 95% of its teachers certified, which is one of the highest percentages in the state of Louisiana.

A brief description of the eight elementary schools that participated in this study will follow. They are separated into four categories including the following:

1) Poor SPS Score – Lower SES

School A: Campus Description

The following description is offered in order to better understand the climate of the elementary schools where teachers were surveyed. This information is a product of information from members of the respective schools, the school's website, the Louisiana Department of Education's website, and the community websites. Due to the confidentiality agreement made with the respondents, I will not use actual names of schools or participants. All of the participating elementary schools were located in Jefferson Parish. This parish was chosen to conduct the study because it is one of the largest school districts in the state.

Additionally, the district is comprised of a number of diverse elementary schools that fit into the four categories created by the two independent variables, all of which are in suburban areas. The schools were accessible, my seventeen years of experience in the district was important in gaining access to the schools.

School A is located on the west bank of the parish. Many apartment complexes and low-income housing projects surround School A. This school is in a small community of 10,763. The median household income is \$27,218, the median age is 35 years old, and 32.7% of the households have children under 18 years old. Fourteen percent of the members of the community have college degrees. The community is 75.16% is White, 20.02% is African-American, and 3.59% is Hispanic. The local community provides very little support for School A.

School A is a Pre-Kindergarten to fifth grade school. This school averages 8% of its students being absent each day. It has a 16:1 pupil to teacher ratio and a declining enrollment of approximately 180 students with 90% being regular education students and 10% in Special

Education. “Regular education reflects students [without disabilities] who [could also be] considered gifted and talented, Section 504, and those with speech or language impairment. Special education includes those students with disabilities and LEP refers to students who speak English as a Second Language” (Tolbert, 2003, p.69). School A’s racial makeup is 98% African-American and 2% White.

In 2007, School A had an SPS score of 38.6 and 95.24% of the students at this school received free or reduced lunch. One hundred percent of the teachers are certified. Eleven of the thirteen available teachers completed the survey. Throughout this section, I will refer to available teachers because many schools still had several vacancies to fill at the time of the survey’s implementation.

#### School B: Campus Description

School B is located on the west bank of the parish. Many apartment complexes and low-income housing projects surround School B. This school is in a small community of 10,897. The median household income is \$28,065, the median age is 36 years old, and 35.7% of the households have children under 18 years old. Thirteen percent of the members of the community have college degrees. The community is 56.32% is White, 35.53% is African-American, and 6.34% is Hispanic. The community has 24.2% of its members living below the poverty line. The local community provides only token support for School B.

School B is a Pre-Kindergarten to fifth grade school. This school averages 6.4% of its students being absent each day. It has a 14:1 pupil to teacher ratio and a declining enrollment of approximately 238 students with 79.9% being regular education students and 21.1% in Special Education. School B’s racial makeup is 93% African-American, 4% White and 2% Hispanic.

In 2007, School B had an SPS score of 48.4 and 95.38% of the students at this school received free or reduced lunch. Fifty-five percent of the teachers are certified. Thirteen of the fifteen available teachers completed the survey.

## 2) High SPS – Lower SES

### School C: Campus Description

School C is located on the east bank of the parish. Many middle-income houses surround School C. This school is in one of Louisiana's largest communities with a population of 70,517. The median household income is \$39,946, the median age is 34 years old, and 36.3% of the households have children under 18 years old. 26.7% of the members of the community have college degrees. The community is 68.12% is White, 22.55% is African-American, and 13.62% is Hispanic. The community has 13.6% of its members living below the poverty line. Many of the local community businesses provide financial support to the school.

School C is a Pre-Kindergarten to fifth grade school. This school averages 4.1% of its students being absent each day. It has a 13:1 pupil to teacher ratio and an enrollment of approximately 168 students with 88% being regular education students and 12% in Special Education. School C's racial makeup is 92.9% African-American, 5.6% White and 1.5% Hispanic.

In 2007, School C had an SPS score of 73.6 and 99.4% of the students at this school received free or reduced lunch. Ninety-one percent of the teachers are certified. All of the twenty-one available teachers completed the survey.

### School D: Campus Description

School D is located on the east bank of the parish. Due to the closeness of School D to School C; the description of the community of School D is the same as School C.

School D is a Pre-Kindergarten to fifth grade school. This school averages 4.3% of its students being absent each day. It has a 14:1 pupil to teacher ratio and an enrollment of approximately 257 students with 88% being regular education students and 12% in Special Education. School D's racial makeup is 28.6% African-American, 44.5% White and 23.7% Hispanic.

In 2007, School D had an SPS score of 99.6 and 80.54% of the students at this school received free or reduced lunch. All of the teachers are certified. Nineteen of the twenty available teachers completed the survey.

### 3) Poor SPS Score – Higher SES

#### School E: Campus Description

School E is located on the west bank of the parish. Many plantation houses and expensive houses surround School B. This school is in a small tight knit community of 1,576. The median household income is \$33,872, the median age is 36 years old, and 33.5% of the households have children under 18 years old. Twenty-six percent of the members of the community have college degrees. The community is 94.54% White, 1.14% African-American, and 1.78% Hispanic. The community has 15.9% of its members living below the poverty line. The local community provides a great deal of volunteer and financial support for School E.

School E is a Pre-Kindergarten to fifth grade school. This school averages 7.2% of its students being absent each day. It has a 16:1 pupil to teacher ratio and a declining enrollment of approximately 519 students with 88% being regular education students and 12% in Special Education. School E's racial makeup is 90.9% White, 3.4% African-American, and 1.4% Hispanic.

In 2007, School E had an SPS score of 68.8 and 64.55% of the students at this school received free or reduced lunch. Ninety percent of the teachers are certified. Seven of the twelve

available teachers completed the survey. Several attempts were made through the lead teacher and the principal to obtain a higher percentage of teachers to complete the survey; however, some teachers were steadfast in not completing the survey. The principal assured me that the teachers that completed the survey were a fair representation of the school's teachers and "usually speak for the faculty".

#### School F: Campus Description

School F is located on the east bank of the parish. Many middle-income houses surround School F. This school is Jefferson Parish's largest community with a population of 146,136. The median household income is \$41,233, the median age is 36 years old, and 25.3% of the households have children under 18 years old. 28.7% of the members of the community have college degrees. The community is 85.53% is White, 6.83% is African-American, and 7.25% is Hispanic. The community has 8.9% of its members living below the poverty line. Many of the local community businesses provide financial support.

School F is a Pre-Kindergarten to fifth grade school. This school averages 4.3% of its students being absent each day. It has a 16:1 pupil to teacher ratio and an enrollment of approximately 487 students with 88% being regular education students and 12% in Special Education. School F's racial makeup is 34.4% African-American, 30.4% White, 27% Hispanic, and 7.8% Asian.

In 2007, School F had an SPS score of 68.8 and 75.36% of the students at this school received free or reduced lunch. All of the teachers are certified. Fourteen of the eighteen available teachers completed the survey.

#### 4) High SPS – Higher SES

##### School G: Campus Description

School G is located on the east bank of the parish. Many upper middle-income houses surround School G. This school is a small community with a population of 9,885. The median household income is \$44,702, the median age is 36 years old, and 34.9% of the households have children under 18 years old. 31.2% of the members of the community have college degrees. The community is 97.06% is White, 0.62% is African-American, and 2.46% is Hispanic. The community has 8.6% of its members living below the poverty line. Many of the local community businesses provide financial support.

School G is a Pre-Kindergarten to fifth grade school. This school averages 4.4% of its students being absent each day. It has a 16:1 pupil to teacher ratio and an enrollment of approximately 460 students with 85% being regular education students and 15% in Special Education. School G's racial makeup is 26.1% African-American, 67% White and 3.4% Hispanic.

In 2007, School G had an SPS score of 93.3 and 55.87% of the students at this school received free or reduced lunch. One hundred percent of the teachers are certified. Nineteen of the twenty-one available teachers completed the survey.

##### School H: Campus Description

School H is located on the east bank of the parish. Due to the closeness of School H to School F; the description of the community of School H is the same as School F.

School H is a Pre-Kindergarten to fifth grade school. This school averages 5.2% of its students being absent each day. It has a 16:1 pupil to teacher ratio and an enrollment of approximately 306 students with 68% being regular education students and 32% in Special



Education. School H's racial makeup is 69.4% White, 30.4% African-American, and 6.4% Hispanic.

In 2007, School H had an SPS score of 94.9 and 48.37% of the students at this school received free or reduced lunch. Ninety-five percent of the teachers are certified. Fourteen of the twenty-one available teachers completed the survey. Seven teachers submitted the surveys with more than one answer in the questions or left questions blank.

#### Results from Phase II Study

A total of 125 surveys were returned out of a possible 141 surveys distributed to teachers in 8 different schools for an 89% return rate. Some schools returned unfilled surveys. These are indicated by non-responses in the table on the next page. One survey at School H contained multiple responses, so this survey was classified as non-response as well. All non-responses were eliminated from the study. As a result, 112 of a possible 141 or 80% of the surveys were utilized for this study. The effective sample size from each school is indicated in Table 4.1 on the next page as well.

There were also some surveys that contained partial responses. These were included in the study and statistical procedures were modified to incorporate missing values. Default missing value procedures in SPSS were used to run unbalanced ANOVAs to accommodate the partial responses. The causal comparative design and sample sizes are detailed below.

#### Independent Variables in the Study

Each survey consisted of 21 fixed response questions and one open ended question. The SPSS data file that I created contains the school name, SPS level (high or poor), Soc/Eco (higher or lower SES), and the respondents' answers to questions 6 – 21. Questions 2, 3, 4, and 5 did not pertain directly to the research questions.

Table 4.1:

Effective Sample Size

School	SPS Level	SES Level	No. of Surveys Administered	No. of Non-responses	Effective Sample Size
School A	Poor	Lower	11	0	11
School B	Poor	Lower	13	0	13
School C	High	Lower	21	0	21
School D	High	Lower	19	0	19
School E	Poor	Higher	7	0	7
School F	Poor	Higher	14	6	8
School G	High	Higher	19	0	19
School H	High	Higher	21	7	14
<b>Totals</b>			<b>125</b>	<b>13</b>	<b>112</b>

Note. 112 of a possible 141 or 80% of the surveys were utilized for this study. SPS=School Performance Score. SES= Socioeconomic Status.

These questions were used as demographic information that was included in the section describing the schools in this chapter. In addition, these questions allowed for more than one response that made recording in a single variable field impossible.

The research questions were used together to create the three dependent variables. Each of these is described below.

Research question one addressed the effect of LEAP testing on classroom practice. In the survey, questions 11 – 15 pertain to classroom practices. A single dependent variable, called “practice” in the SPSS data file, was created by summing the responses to these questions and the new variable was used as a measure of the perceived effect LEAP testing has on classroom practice. Note that each of the survey questions was worded in a way that a 1 indicated low

effect and 5 indicated a high effect. Thus, for each respondent, the variable “practice” can take on values 5 (low overall effect) to 25 (high overall effect).

Research question two addressed the perceived pressure to spend more time on test preparation. In the survey, questions 16 – 19 pertain to this perceived pressure. Question 19 was subdivided into 6 questions, so there were essentially 9 survey questions that measured perceived pressure. A single dependent variable, called “pressure” in the SPSS data file, was created by summing the responses to these 9 questions and the new variable was used as a measure of the perceived pressure to spend more time on test preparation. Note that each of the survey questions was worded in a way that a 1 indicated low effect and 5 indicated a high effect. Thus, for each respondent, the variable “pressure” can take on values 9 (low overall effect) to 45 (high overall effect).

Research question three addressed the degree of commitment to the profession among the respondents. In the survey, questions 7 – 9 pertain to the respondent’s degree of commitment. A single dependent variable, called “commit” in the SPSS data file, was created by summing the responses to these questions and the new variable was used as a measure of each respondent’s commitment to the profession. Note that questions 7, 8, and 9 were worded in a way that a 1 indicated a high commitment and 5 indicated a low commitment. Thus, for each respondent, the variable “commit” can take on values 3 (high overall commitment) to 15 (low overall commitment).

### Descriptive Statistics for Independent Variables

Each of the dependent variables described above have different sample sizes due to partial responses. I address the sample sizes and provide descriptive statistics for each variable in the forthcoming sections. In the tables,  $\bar{x}$  stands for the sample mean and  $s$  stands for the sample

standard deviation. I provide a breakdown of the descriptive statistics by the levels of the independent variable. This makes it easier to explain the results of the analyses later.

Classroom Practice Variable

Of the 112 surveys used in the analysis, 9 respondents gave partial responses to questions related to perceived effect on classroom practices. Thus, there were only 103 data values available for this variable. A breakdown of these responses is contained in Table 4.2. This table displays the amount of surveys that were utilized concerning this variable separated by the four cells created during the sampling phase of this study.

Table 4.2

Responses Concerning Classroom Practices

		SES	
SPS		Higher SES	Lower SES
Poor		15	24
High		33	40

Note. Higher SES = 75.0 or greater, Lower SES = 74.9 or lower, and SPS High = 73.0 or greater, SPS Poor = 72.9 or lower. SPS=School Performance Score. SES= Socioeconomic Status.

The means and standard deviations for the “practice” variable are given in Table 4.3. Marginal means are included to assist in comparisons after ANOVAs are run. Questions 11 – 15 in the survey that was utilized for this study were devoted to the dependent variable, classroom practices. A teacher that completed the survey would have encountered a Likert scale that produced questions with five choices. Those choices ranged from “1 = Strongly Disagree” to “5= Strongly Agree”.

Circling “1” meant that your daily classroom practices were unaffected by LEAP 21 testing. On the other hand, circling “5” meant that your daily classroom practices were greatly affected by LEAP 21 testing.

The questions were combined to address the first research hypothesis. The scores of the means ranged from “5” (low effect) to “25” (high effect) with a midpoint of “15”. All the cells’ mean scores were above average and denote that regardless of high or low SPS scores or higher or lower SES status, all respondents perceived their classroom practices to be affected by LEAP 21 testing.

Table 4.3:

Means and Standard Deviations Concerning Classroom Practices

SPS	SES		Marginal Means
	Higher SES	Lower SES	
Poor	$\bar{x} = 18.93, s = 2.43$	$\bar{x} = 16.38, s = 2.73$	17.44
High	$\bar{x} = 18.31, s = 3.43$	$\bar{x} = 18.29, s = 2.86$	18.30
Marginal Means	18.52	17.61	

Note. The scores of the means ranged from “5” (low effect) to “25” (high effect) with a midpoint of “15”. All the cells’ mean scores were above average and denote that regardless of high or low SPS scores or higher or lower SES status, all respondents perceived their classroom practices to be affected by LEAP 21 testing. SPS=School Performance Score. SES= Socioeconomic Status.

Perceived Pressure Variable

Of the 112 surveys used in the analysis, 14 respondents gave partial responses to questions related to perceived pressure. Thus, there were only 98 data values available for this variable. A breakdown of these responses is contained in Table 4.4.

Table 4.4:

Responses Concerning Perceived Pressure

SPS	SES	
	Higher SES	Lower SES
Poor	15	20
High	28	35

Note. Higher SES = 75.0 or greater, Lower SES = 74.9 or lower, and SPS High = 73.0 or greater, SPS Poor = 72.9 or lower. SPS=School Performance Score. SES= Socioeconomic Status.

The means and standard deviations for the “pressure” variable are given in Table 4.5 below. Marginal means are included to assist in comparisons after ANOVAs are run.

Table 4.5:

Means and Standard Deviations Concerning Perceived Pressure

SPS	SES		Marginal Means
	Higher SES	Lower SES	
Poor	$\bar{x} = 27.47, s = 6.48$	$\bar{x} = 28.25, s = 5.52$	27.91
High	$\bar{x} = 27.79, s = 6.70$	$\bar{x} = 29.46, s = 6.46$	28.71
Marginal Means	27.68	29.02	

Note. The scores of the means ranged from “9” (low effect) to “45” (high effect) with a midpoint of “27”. All the cells’ mean scores were above average and denote that regardless of high or low SPS scores or higher or lower SES status, all respondents perceived that pressure was present due to LEAP 21 testing. SPS=School Performance Score. SES= Socioeconomic Status.

Questions 16 – 19 in the survey that was utilized for this study were devoted to the dependent variable, pressure. Question 19 was divided into six categories. A teacher that

completed the survey would have encountered a Likert scale that produced questions with five choices. Those choices ranged from “1 = Strongly Disagree” to “5= Strongly Agree”.

Circling “1” meant that these teachers perceived little to no pressure regarding LEAP 21 testing. On the other hand, circling “5” meant that these teachers perceived a great amount of pressure regarding LEAP 21 testing.

The questions were combined to address the second research hypothesis. The scores of the means ranged from “9” (low pressure) to “45” (high pressure) with a midpoint of “27”. All the cells’ mean scores were above average and denote that regardless of high or low SPS scores or higher or lower SES status, all respondents perceived that pressure was present due to LEAP 21 testing.

Degree of Commitment Variable

Of the 112 surveys used in the analysis, 3 respondents gave partial responses to questions related to degree of commitment. Thus, there were only 109 data values available for this variable. A breakdown of these responses is contained in Table 4.6.

Table 4.6  
Responses Concerning Commitment

SPS	SES	
	Higher SES	Lower SES
Poor	15	23
High	32	39

Note. Higher SES = 75.0 or greater, Lower SES = 74.9 or lower, and SPS High = 73.0 or greater, SPS Poor = 72.9 or lower. SPS=School Performance Score. SES= Socioeconomic Status.

The means and standard deviations for the “commit” variable are given in Table 4.7 below. Marginal means are included to assist in comparisons after ANOVAs are run.

Questions 7 – 9 in the survey that was utilized for this study were devoted to the dependent variable, commitment. A teacher that completed the survey would have encountered a Likert scale that produced questions with five choices. Those choices ranged from “1 = Strongly Disagree” to “5= Strongly Agree”.

Unlike the previous questions, circling “1” meant that these teachers perceived that they had remained committed in light of the LEAP 21 testing. On the other hand, circling “5” meant that these teachers perceived a decreased commitment due to LEAP 21 testing.

The questions were combined to address the third research hypothesis. The scores of the means ranged from “3” (high commitment) to “15” (low commitment) with a midpoint of “9”. All the cells’ mean scores were slightly below average and denote that regardless of high or low SPS scores or higher or lower SES status, all the respondents on average perceived that LEAP 21 testing did not have an effect on the teachers’ degree of commitment.

Table 4.7:  
Means and Standard Deviations Concerning Commitment

SPS	SES		Marginal Means
	Higher SES	Lower SES	
Poor	$\bar{x} = 8.33, s = 1.11$	$\bar{x} = 8.48, s = 1.95$	8.41
High	$\bar{x} = 7.88, s = 1.79$	$\bar{x} = 8.48, s = 1.62$	8.21
Marginal Means	8.02	8.48	

Note. The scores of the means ranged from “3” (high commitment) to “15” (low commitment) with a midpoint of “9”. All the cells’ mean scores were slightly below average and denote that regardless of high or low SPS scores or higher or lower SES status, all the respondents on average perceived that LEAP 21 testing did not have an effect on the teachers’ degree of commitment.



## Analysis of Research Hypotheses

### Rationale for Analysis

Throughout this section there is a series of two-step analysis:

- (1) First, one-way ANOVAs are reported (with poor or high SPS scores as the independent variables) that test each of the three hypotheses, each of which require such an analysis.
- (2) Second, two-way ANOVAs are reported (with poor or high SPS scores and lower or higher SES as independent variables) to assess the effect that these variables jointly have on the dependent variables.

The first analyses test the hypotheses, while the second analyses provide more information on the context within which the accountability program exists.

### Research Question 1

Paraphrasing, the hypothesis was that teachers' perception of the LEAP testing effect on classroom practices would depend on the school's SPS scores (poor vs. high). To test this hypothesis, a one-way ANOVA was run with the variable "classroom practice," defined previously, as the dependent variable and SPS as the independent variable. The ANOVA table is on page 95.

Utilizing Table 4.8, I sought to determine whether or not the mean classroom practice variable was different for schools experiencing past testing success or failure (SPS) in regards to the classroom practice variable. The significance level associated with SPS is .174 which is not statistically significant thereby indicating that there is not a difference between the low and high scores with regard to the classroom practice variable.

In regards to the first research hypothesis, the ANOVA table indicates no significant difference in perceived effect on classroom practices between schools scoring high and schools

scoring poor on the LEAP 21 test (SPS score). From Table 4.3 of this study, we see that schools scoring higher on the LEAP 21 test had higher mean scores on classroom practices. Higher scores meant that their classroom practices were perceived to be more affected by LEAP 21 testing. However, this difference is not statistically significant. This means that schools which experienced past testing success did not perceive a significantly greater effect or change to their classroom practices. This finding fails to confirm the hypothesis.

Data were also collected on the socio-economic make-up of the schools as measured by the number of students on free/reduced lunch. The sampled schools were divided into two categories: higher SES and lower SES. To test the hypothesis while jointly testing for differences among the socio-economic levels, a two-way ANOVA was run with the variable “practice” as the dependent variable and SPS and SES as independent variables. The purpose of utilizing the two-way ANOVA was to discover if there was a difference in classroom practice in regard to SPS (high and low) while taking into consideration the socioeconomic status (SES) of the schools. The ANOVA table (Table 4.9) is on page 96.

The Table 4.9 indicates a significant interaction between the two independent variables meaning that the difference in response between SPS high and SPS poor depends upon the socio-economic level of the school. This is again evident in Table 4.3 of this report. Note that in this table, although all of the categories scored high, Poor SPS/ Lower SES cell reported somewhat lower on classroom practices; whereas, the rest of the categories reported similarly high scores. This one cell accounts for the interaction.

Bruce Thompson (1998, p. 34), an outspoken advocate of researchers including effect size in their studies, has stated:

More than anything else, I especially want to see authors always report effect sizes. I concur with the views of McLean and Ernest (1998), who noted that, ‘In reviewing the literature, the authors were unable to find an article that argued against the value of including some form of effect size or practical significance estimate in a research report.

Table 4.8:

Tests of Between-Subjects Effects (SPS) Dependent Variable: Practice

<i>Source</i>	<i>Type III SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>p</i>
SPS	17.081	1	17.081	1.873	.174
Error	920.919	101	9.118		
Total	34310.000	103			
Corrected Total	938.000	102			

Note. In regards to the first research hypothesis, the ANOVA table indicates that there is not a significant effect for the SPS independent variable on the practice dependent variable.

To present a measure of the substantive significance of the difference between these measures, the partial eta squared ( $\eta_p^2$ ) for the interaction was calculated as an effect size and was found to be .041 which is interpreted to be a small effect size. Nonetheless, the interaction was significant indicating that the effect of SPS on classroom practice depends on the socioeconomic status of the students. Since there is a significant interaction, the main effects SPS and SES cannot be interpreted from this two-way ANOVA.

Research Question 2

Paraphrasing, the hypothesis was that teachers' perception of LEAP testing creating pressure to spend more time on test preparation would depend on the school's SPS scores (poor vs. high). To test this hypothesis, a one-way ANOVA was run with the variable "pressure" defined previously, as the dependent variable and SPS as the independent variable.

Table 4.9:

Tests of Between-Subjects Effects (SPS\*SES) Dependent Variable: Practice

<i>Source</i>	<i>Type III SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>P</i>
SPS	17.081	1	17.081	1.957	.165
SES	20.408	1	20.408	2.339	.129
SPS * SES	36.603	1	36.603	4.194	.043
Error	863.908	99	8.726		
Total	34310.000	103			
Corrected Total	938.000	102			

Note. The interaction was significant indicating that the effect of SPS on classroom practice depends on the socioeconomic status of the students.

Utilizing Table 4.10, I sought to determine whether or not the mean pressure variable was different for schools experiencing past testing success or failure (SPS) in regards to the pressure variable. The significance level associated with SPS is .550 indicating that there is not a significant difference between the low and high scores with regard to the pressure variable.

In regards to the second research hypothesis, the ANOVA table indicates that there is not a significant effect for the SPS independent variable on the dependent variable pressure. From Table 4.5 of this study, we see that schools scoring higher indicated a higher perceived pressure to spend more time on test preparation. Higher mean scores meant that the amount of pressure was perceived to be accelerated by LEAP 21 testing. However, this difference is not statistically significant. This means that schools which experienced past testing success did not perceive a significantly greater effect or change in the amount of pressure experienced. This finding fails to confirm the hypothesis.

Table 4.10:

Tests of Between-Subjects Effects (SPS) Dependent Variable: Pressure

<i>Source</i>	<i>Type III SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>p</i>
SPS	14.400	1	14.400	.359	.550
Error	3847.600	96	40.079		
Total	83064.000	98			
Corrected Total	3862.000	97			

Note. In regards to the second research hypothesis, the ANOVA table indicates that there is not a significant effect for the SPS independent variable on the pressure dependent variable.

To test the hypothesis while jointly testing for differences among the socio-economic levels, a two-way ANOVA was run with the variable “pressure” as the dependent variable and SPS and SES as independent variables. The ANOVA table is below.

Table 4.11:

Tests of Between-Subjects Effects (SPS\*SES) Dependent Variable: Pressure

<i>Source</i>	<i>Type III SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>p</i>
SPS	14.40	1	14.40	.356	.552
SES	44.36	1	44.36	1.098	.297
SPS * SES	4.36	1	4.36	.108	.743
Error	3798.88	94	40.41		
Total	83064.00	98			
Corrected Total	3862.00	97			

Note. The interaction was not significant indicating that the effect of SPS on pressure does not depend on the socioeconomic status of the students.

Table 4.11 indicates that no interaction effect exists. The interaction test tells you whether the effects of one factor depend on the other factor. The main effects of SPS and SES as

demonstrated in the two-way ANOVA table indicated no significant difference in perceived pressure between schools scoring high and schools scoring poor on the SPS or schools with high or lower SES.

The results of these ANOVAs would seem to indicate that SPS and SES or a combination of the two have little effect on the perceived pressure variable. The hypothesis is not only unproven but because of Table 4.5, there is evidence that the opposite is true.

Research Question 3

Paraphrasing, the hypothesis was that teachers’ degree of commitment would depend on the school’s SPS scores (poor vs. high). To test this hypothesis, a one-way ANOVA was run with the variable “commit,” defined previously, as the dependent variable and SPS as the independent variable. The ANOVA table is below.

Table 4.12:

Tests of Between-Subjects Effects (SPS) Dependent Variable: Commit

<i>Source</i>	<i>Type III SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>p</i>
SPS	1.09	1	1.09	0.38	.539
Error	307.09	107	2.87		
Total	7789.00	109			
Corrected Total	308.18	108			

Note. In regards to the third research hypothesis, the ANOVA table indicates that there is not a significant effect for the SPS independent variable on the commit dependent variable.

The ANOVA table indicates no significant difference in the degree of commitment between schools scoring high and schools scoring poor on the SPS. Again, the significance level

associated with SPS is .539 which is not statistically significant indicating that there is not a difference between the low and high scores with regard to the commit variable.

In regards to the third research hypothesis, the ANOVA table indicates that there is not a significant effect for the SPS independent variable on the commit dependent variable. From Table 4.7 of this study, we see that schools scoring lower indicated a higher perceived commitment to the education profession. Lower scores meant that the amount of commitment was affected by LEAP 21 testing. However, this difference in mean scores was not statistically significant. This means that schools which experienced past testing success did not perceive a significantly greater amount of commitment to the educational profession. This finding fails to confirm the hypothesis.

To test the hypothesis while jointly testing for differences among the socio-economic levels, a two-way was run with the variable “pressure” as the dependent variable and SPS and SES as independent variables. The ANOVA table (Table 4.13) is on the next page.

Table 4.13 indicates that no interaction effect exists. The interaction test tells you whether the effects of one factor depend on the other factor. The main effects of SPS and SES as demonstrated in the two-way ANOVA table indicated no significant difference in commitment between schools scoring high and schools scoring poor on the SPS or schools with high or lower SES.

The results of these ANOVAs would seem to indicate that SPS and SES or a combination of the two have little effect on the commitment variable. The hypothesis was not confirmed.

#### Summary of Chapter 4

The quantitative results of this research were presented in the present chapter. A sampling of eight elementary schools was taken from the population of Jefferson Parish Public Schools by

utilizing a stratified purposive sampling technique. In this chapter the Research Hypotheses were reintroduced and quantitative data collection techniques (surveys) were described.

A brief description of the eight elementary schools and their surrounding communities that participated in this study was presented. The surveys that were returned were then analyzed and some were eliminated for the reasons stated.

Each survey consisted of 21 fixed response questions and one open ended question. The SPSS data file that was created contained the school name, SPS level (high or poor), SES (higher or lower SES numbers), and the respondents' answers to the questions. The survey questions were grouped together to address the three dependent variables: classroom practices, pressure, and commitment. Means, standard deviations, and one and two-way ANOVAs were utilized to calculate the data.

Table 4.13:

Tests of Between-Subjects Effects (SPS\*SES) Dependent Variable: Commit

<i>Source</i>	<i>Type III SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>p</i>
SPS	1.089	1	1.089	0.38	.539
SES	5.471	1	5.471	1.91	.169
SPS * SES	1.307	1	1.307	0.46	.499
Error	300.316	105	2.860		
Total	12201.000	109			
Corrected Total	308.183	108			

Note. The interaction was not significant indicating that the effect of SPS on commitment does not depend on the socioeconomic status of the students.

The first hypothesis was that teachers' perception of the LEAP testing effect on classroom practices would depend on the school's SPS scores (poor vs. high). A Two-Way



ANOVA found that a significant interaction between the two independent variables meaning that the difference in response between SPS high and SPS poor depends upon the socio-economic level of the school. This was also evident through the means and standard deviations collected for this report. Although all of the categories scored high, the Poor SPS/ Lower SES cell reported less of an effect on classroom practices; whereas, the rest of the categories reported similar high scores. This one cell accounts for the interaction.

The second hypothesis was that teachers' perception of LEAP testing creating pressure to spend more time on test preparation would depend on the school's SPS scores (poor vs. high). Surprisingly, Table 4.5 demonstrated that schools scoring higher indicated a higher perceived pressure to spend more time on test preparation. In fact, the category that accounted for the most pressure was High SPS/ Lower SES.

The third hypothesis was that teachers' degree of commitment would depend on the school's SPS scores (poor vs. high). An ANOVA indicated a significant difference does not exist in the degree of commitment between schools scoring high and schools scoring poor on the SPS.

The following four points summarize the results of the quantitative analyses:

- (1) How well students performed on the high stakes testing and the SES of students at the schools had little effect on their teachers' perceptions and responses to the testing program.
- (2) This is surprising since there are reasons to believe that teachers at high performing schools would perceive and react differently to high stakes testing. This was not the case, except for perceived pressure, where the effect was opposite to prediction.
- (3) The teachers' overall scores on the practice (around 18) and pressure (around 28) variables were above the midpoint for those scales (15 and 27 respectively), indicating relatively high overall levels for all teachers on these variables.

(4) The teachers' scores on the commitment variable was around 8, which is slightly below the midpoint of that scale, indicating a relatively slight commitment overall on the part of all the teachers.

To portray completely the impact of LEAP/iLEAP testing, a qualitative analysis was needed.

The results of the qualitative phase (Phase III) of the study are presented in the next chapter.

## CHAPTER 5: QUALITATIVE RESULTS

### Introduction

The people who have the most knowledge of the effects of LEAP/iLEAP testing are those who are closest to the testing — the teachers. However, the teachers' voices are often ignored in the testing debate. “Their views are often dismissed by testing advocates who argue that teachers oppose high stakes tests simply because they do not want to be held accountable; teachers are biased so their concerns about high stakes tests should not be warranted” (Wright, 2002, p. 5). Yanow (2000), however, states, "To understand the consequences of a policy for the broad range of people it will affect requires 'local knowledge'- the very mundane, expert understandings of and practical reasoning about local conditions derived from lived experience" (pp. 4-5). Without question, teachers are the guardians of this "local knowledge" and have the “experience” necessary to understand the consequences of high-stakes testing.

In fact, it may be harmful to any educational change to ignore the voice of the teachers. As Salvaterra & Adams stated "... any major change, may hinge more on teachers' perceptions of the change than on its actual merits" (1995, p. 35). I chose qualitative interview methods using the interpretist theory (Erickson, 1986) as a guide. Interpretist researchers accept as true that the most truthful means of understanding phenomena is to consider insider viewpoints (Eisenhart & Howe, 1990). In this study, first person accounts of testing were necessary to supply a basis for inferences about perceived implications. Therefore, qualitative research methodologies (interviews) were employed to accumulate the data needed to discover this "local knowledge" (Bogdan & Biklen, 1998).

In this chapter, a brief description of the teachers who participated in the study is presented. Due to the promise of confidentiality, the identities of the participants of this study are

not disclosed. The basic demographic information provided by the respondents is included in this chapter.

A sampling of 16 elementary school teachers interviewed was taken from the population of eight Jefferson Parish Public Schools that had been surveyed. A stratified purposive sampling technique, described in detail in Chapter 3, divided the purposefully selected target population into the following strata: elementary schools that earned high SPS scores and elementary schools that had low SPS scores. These groups were split once again into the following strata: elementary schools that have a high percentage of students in the free and reduced lunch program and those that have a smaller percentage. As explained before, researchers often use the free/reduced lunch program as an indicator of students from poor socioeconomic backgrounds. These actions created four groupings. Two schools were chosen from each of the four groupings creating a total of eight schools utilized in this report. Two elementary teachers from each of the eight schools were purposely selected for this study.

The criteria for choosing the teachers to interview are described in Chapter 3. The interviews were conducted over a week period in the 2007-2008 fall semester before testing was conducted. The interview questions were informed by the survey responses. The answers to the Research Hypotheses that guided the surveys are summarized in Table 5.2. The questions that helped guide the interviews can be found in Appendix D.

As stated earlier the qualitative portion of this study will address the research questions. The research questions are as follows:

#### Research Questions

- 1 How does test preparation (teaching the test) affect teachers' instructional planning, learning strategies, and curriculum content and to what extent?

- 2 How much time do teachers perceive that students spend on test preparation and how does that amount of time compare to the time spent on instruction?
- 3 What effect does testing have on an educators' sense of professionalism and pride in their work? How does high stakes testing affect motivation in general?

### Participants

The results of the study are based on responses from elementary school teachers from the Jefferson Parish Public School System. The sample of teachers utilized for the study was the result of a phased selection process. First, eight elementary schools were selected utilizing the criteria set forth in Chapter 3. Identified schools were then contacted, and all elementary school teachers within them asked to complete a survey. Upon completion of the surveys it was determined that those teachers who were most knowledgeable about high stakes testing were teachers from grades 3, 4, and 5. This makes perfect sense since these are the grades tested at the elementary school level.

This is an example of intensity sampling (Patton, 2002).

Intensity sampling involves the same logic as extreme case sampling but with less emphasis on the extremes. An intensity sample consists of information-rich cases that manifest the phenomenon of interest intensely (but not extremely). Extreme or deviant cases may be so unusual as to distort the manifestation of the phenomenon of interest. Using the logic of intensity sampling, one seeks excellent or rich examples of the phenomenon of interest, but not highly unusual cases (p. 234).

Armed with the information from the surveys, I felt it was important to interview teachers who were both experienced and knowledgeable about LEAP/iLEAP tests. I contacted the principals of each school and asked them to recommend two teachers from grades 3, 4, and 5 or the LEAP Coordinator and one teacher from grades 3, 4, and 5. All but two schools chose the latter. This resulted in two kindergarten teachers and one second grade teacher being interviewed. The rest of the teachers interviewed were from grades 3, 4, and 5 with the majority being from 4<sup>th</sup> grade. A modest honorarium was offered to each teacher participating in the

interviews. The overwhelming majority of the teachers interviewed had extensive teaching experience with only three of the sixteen having less than six years experience. The average number of years of experience of the interview participants was 13 years teaching in Jefferson Parish Public Schools. Ten of the participants were Caucasian and six were African- American. All of the interviewees except one were female. The promise of anonymity was maintained by referring to the individual interviewees by school letter and whether they were the first teacher interviewed or the second. For instance, the first person I interviewed at School C is referred to as “C-1”. Table 5.1 has information on the teachers interviewed.

#### Data Collection Procedures

After the principals recommended the teachers to be interviewed, I then proceeded to contact each participant by phone at their respective schools. I explained to them the purpose of the study. When they agreed to an interview, an appointment was scheduled. All teachers so contacted agreed to participate in this study. Next, principals and interview participants were sent a letter that gave a brief abstract of the study; the procedures for the interview; and their rights regarding the interview (Appendix F). A small honorarium was presented to each participant at the end of their interview. The interviews were conducted at the respective school sites during the second week of November.

To conduct these interviews an interview guide approach was utilized. The interview guide (Patton, 2002) that was created can be found in Appendix D. This list of questions provided me with the flexibility to explore further concepts during each interview. Although it was prepared to insure that basically the same information was obtained from each person, there was no predetermined responses. I was free to probe and explore within these predetermined topics. Interviews were taped and transcribed shortly thereafter.

Table 5.1:

Teachers Interviewed

Teacher	School's SPS Score	School's SES	Years of Experience	Grade Taught	Gender
A-1	Poor	Lower	30	4	F
A-2	Poor	Lower	3	3	F
B-1	Poor	Lower	3	3	F
B-2	Poor	Lower	12	4	F
C-1	High	Lower	30	4	F
C-2	High	Lower	6	4	F
D-1	High	Lower	12	4	M
D-2	High	Lower	5	4	F
E-1	Poor	Higher	11	4	F
E-2	Poor	Higher	10	K	F
F-1	Poor	Higher	8	2	F
F-2	Poor	Higher	11	K	F
G-1	High	Higher	18	5	F
G-2	High	Higher	23	5	F
H-1	High	Higher	17	4	F
H-2	High	Higher	9	4	F

Note. K = Kindergarten.

Data Analysis Procedures

Originally, the strategy employed was to split up the 16 interviews into the same four cells that were created during the quantitative phase of this study and to analyze the groups separately looking for different patterns and themes. Four teachers were each assigned to High

SPS/Lower SES, Poor SPS/Lower SES, High SPS/Higher SES, and High SPS/Lower SES.

However, not unlike the quantitative results, there was little difference detected between the four cells as patterns and themes were emerging among the interview data. This occurrence led to a change of strategy. “Design flexibility stems from the open-ended nature of naturalistic inquiry as well as pragmatic considerations” (Patton, 1990, p. 44).

This is not uncommon according to Hoepfl (1997, p.52) “Qualitative research has an emergent (as opposed to predetermined) design, and researchers focus on this emerging process as well as the outcomes or product of the research... It is important to emphasize the emergent nature of qualitative research design. Because the researcher seeks to observe and interpret meanings in context, it is neither possible nor appropriate to finalize research strategies before data collection has begun (Patton, 1990)”. It was therefore decided, after consultation with the peer debriefers, that these interviewees not be separated by groups to be contrasted, but instead to be grouped together and identified in order to show the solidarity, breath, and range of the teachers’ perceptions.

The transcripts that were created by the interviews were read and reread. Once the material was familiar and patterns and themes began to emerge, codes and categories were created. These codes and categories were then entered into a program utilizing the “Atlas.ti” software. This assisted in recognizing and organizing more patterns of information. This data analysis can best be described as the constant comparative method (Glaser, 1978). The exact steps that were followed are chronicled in Chapter 3.

Searching for the appropriate coding, patterns, categories, and themes were constantly employed to reduce the data into manageable entities. Three inclusive categories emerged from the interviews. These categories included: instruction, pressure, and commitment. After the data had been coded and organized into categories, the next step was to generate meaning across the



interviews. I also explored how this data addressed the research questions. Taking the data record as a whole, I progressed inductively to generate assertions.

Within the categories emerged subset categories or themes. Under “Instruction” the following themes surfaced: Teaching to the Test, Neglecting Subjects, Time, Fairness, and Focus on Instruction. Under the category of “Pressure” two themes emerged: Student and Teacher Pressure. No subset themes were discovered under “Commitment”. The result was a total of eight themes to be addressed in this chapter.

Member checks and expert checks of the preliminary research were also conducted. The interview data produced overall themes that included instruction, pressure, and commitment. The next section will address the first research question.

### Instruction

The teachers reported that the Parish and State's emphasis on the LEAP/iLEAP has resulted in drastic changes to the curriculum. In the majority of these cases the teachers did not consider these changes to be positive. Specifically, they stated that in creating their lesson plans they look at prior LEAP tests to make sure that their curricula includes all or most of the test content and assure that they cover test objectives. Further, they report that the test forced them to focus on breadth more than on depth of coverage of the material. They also adjust the sequence of their curriculum based on what is included in the test. In this section the themes of teaching to the test, neglecting subjects, time, fairness, and focus on instruction will be addressed.

1. How does test preparation (teaching the test) affect teachers’ instructional planning, learning strategies, and curriculum content and to what extent?

### Teaching to the Test

The elementary teachers reported that preparing for the LEAP had changed their instructional and assessment strategies and classroom practices. They reported a greater

emphasis on content that would be tested. In addition to these changes in what was taught, most interviewees felt that preparation for the LEAP/iLEAP test had produced changes in how they taught.

It doesn't allow for as much creativity because you have to rigidly teach the test. I wanted to do a project on writing by going through advertisements (other than LEAP) but the principal said that it is usually descriptive and expository which isn't tested ... Practically all of it is directed to the LEAP tests (Teacher A-1, from poor SPS/lower SES school).

Everything you do, you have to think about the test the whole time (Teacher B-1, from poor SPS/lower SES school).

... [Y]ou teach to the standards that the State has given you. The GLEs - so there's a change from just going through the textbook to just addressing the GLEs. I think people do stress the content of the test ... [I teach] test taking skills. I say it starts in January, you see all those Soar books and those test prep books. I say from January to March [we teach] test prep skills - I'd say 70% [of the time] ... I feel like a factory worker I'm just handing out stuff. It's an assembly line. I feel I have no control over what I think is beneficial for the kids ... You don't get a chance to think through. We don't go deep enough. We have to know the nine planets. We just cover the GLEs. It is a detriment; they say it doesn't but it does. You have to work on these skills. It's too broad - all the strands of social studies. They have to know in depth everything. They need to map the curriculum a little bit better (Teacher C-1, from high SPS/lower SES school).

... [In January] we shut down and do nothing but prep for the test. I also prep pre-January but when January comes its pretty much I would say 1/3 of the day -exactly devoted to practicing the test (Teacher C-2, from high SPS/lower SES school).

I gear most of my lessons toward exactly what I feel that is pertinent to LEAP ... I give them tests that are designed in the [LEAP] format (Teacher D-1, from high SPS/lower SES school).

Everything's geared towards the LEAP. I'm confused what strategies I'm using? I may do a different lesson - you know different topics. I could broaden the student's horizons but all I'm focusing on is what's on the LEAP test (Teacher D-2, from high SPS/lower SES school).

We do everything headed to the LEAP. All of our questions are multiple choice and constructed response. All of our assignments are not just fill-in-the blank. It's write a sentence - write a paragraph. I also teach social studies which is on the LEAP but doesn't count, but we still do constructed responses - LEAP prep test questions. Everything is LEAP (Teacher E-1, from poor SPS/ higher SES school).

[We are] teaching for the tests. It's not like when I used to teach units or themes that children like, I feel I can't teach the basic skills, I'm teaching more of what is covered on

the tests and if they don't get [understand] the materials ... [teacher shrugged shoulders] (Teacher G-2, from high SPS/ higher SES school).

I'm always thinking of how can I make this towards the test. I don't know the questions on the test but I know how they ask the questions. I'm always thinking of how are they going to present this on the test (Teacher H-1, from high SPS/ higher SES school).

I'm always thinking in the back of my mind, how am I going to present this as the LEAP test? Once it is January, I would say at least 65% of my day [is devoted to test preparation]. Remember when you see something like this on the test this is what you have to think about or remember to bubble in this way. Let's look at this question - this is how they ask it on the test, 65% - from January to March (Teacher H-2, from high SPS/ higher SES school).

One teacher conceded that she taught to the test. However, she did not see teaching to the test as a problem.

When I first started in 4<sup>th</sup> grade, I was totally against it [LEAP testing] until I got my scores back and I wondered why they didn't learn it. I was teaching the 4<sup>th</sup> grade curriculum but then I got on the band wagon and taught the test. If they learn that and they get it for the test then they do become better life long test takers. They benefited more when I went strictly to the LEAP tests. I still use the test book and follow the curriculum but in a totally different way (Teacher E-2, from poor SPS/ higher SES school).

The majority of teachers felt that not only was the content and focus of their curriculum greatly affected, but that their control of the sequencing of the lessons had changed. This was looked at as a detriment to student learning.

Because the test is given in March, I feel like our time line has to be really pushed close. If there is any time for field trips that are not pertaining to the LEAP, it has to be done in April and May. If there's any kind of unit that might not be covered on LEAP or any kind of just reading for pleasure that may not be a covered topic that are on LEAP, I feel like I have to wait until April or May (Teacher D-1, from high SPS/lower SES school).

I would teach in a different sequence ... by March I know what s expected. I feel like the kids are down here and I feel like what I'm doing is what I'm expected to do. I do get fearful that some of them just aren't going to be able to get it by then (Teacher D-2, from high SPS/lower SES school).

I'm focused on more specific skills and more on detail on specific skills ... I focus on what they tell us to and push all of the other stuff off to the last nine weeks ... focus on certain skills - it gets shelved to the fourth nine weeks like placements, stop at millionth, and shelve like roman numerals and beyond a millionth to the fourth nine weeks (Teacher E-1, from poor SPS/ higher SES school).

Some teachers didn't change the sequence. They were forced to cover some subjects in a less comprehensive manner.

You know what you have to go over but you know it won't be there [on the test] but you still do, but not as strongly as you would normally do. The integral assessment pacing is fast. It is much faster than I would want it to be, but with a small class this year, it isn't too bad but last year it was awful (Teacher B-1, from poor SPS/lower SES school).

### Neglecting Subjects

All of the teachers interviewed, whether they felt that they were coerced or not, discussed that there was a strong propensity among teachers to spend most of their available time (including what little discretionary time they had) on math computation, sight word recognition, comprehension items, and recognition of errors in spelling, usage, punctuation and other basic skills. Reading genuine books, writing in authentic contexts, solving problems that require more than rote recognition, imaginative, critical, and divergent thinking, interactive projects and the like were gradually being replaced in everyday instruction. This trend was the joint consequence of limited time, a crowded curriculum, and the imposition of LEAP/iLEAP testing.

Its everyday - people say it's on the test so we have to cover it. I find there's no enrichment no art or music, I used to bring in songs that you know like *Erie Canal* and *Battle of New Orleans* there doesn't seem to be any time for that any more. It's just cover the curriculum, cover the curriculum. I think that is very much to the detriment of the kids. I think you know we need art in school. We need music. These kids, all kids, love to sing and dance and draw and paint to express themselves. And they have no outlet anymore ... it's made it very cut and dry - cover the curriculum, stand up and lecture, and kids passively are supposed to soak it in - like test taking skills (Teacher C-1, from high SPS/lower SES school).

I have to focus on so much that I can't do a lot of the fun things or I have to shelve it. There's more I would like to delve into but I have to do the LEAP skills and then save it. For instance one of the mystery novels we like to do, we can't do it because mysteries are not on the LEAP test ... I've seen it help a lot of children and hurt a lot of children (Teacher E-1, from poor SPS/ higher SES school).

I've cut out English repetition and circle the noun. What is the noun - who cares? Now we write sentences using the noun and I use like garbage can words instead of just saying "the girl". 98% of the time is devoted to preparing for the tests (Teacher E-2, from poor SPS/ higher SES school).

I find a lot that I don't get to teach as much as I would like because we have these testing things. Well, like I do a lot of art things with my kids and I feel like I can't do a lot of that because we have to do math probes... They need to know other stuff too. I find it takes away from [instruction] in that it isn't as fun as it used to be (Teacher F-1, from poor SPS/ higher SES school).

... [Y]ou got to set aside a time to go over things. You don't have as much time to do group work and [learning] centers and those types of things (Teacher F-2, from poor SPS/ higher SES school).

I think we teach more in the format of the LEAP and iLeap. [There is] material children would benefit from but we needed to leave it out because it is not covered on the test (Teacher G-1, from high SPS/ higher SES school).

It's a hindrance because I feel all we are doing is teaching for the test and not giving them room to explore, because that's not going to be on the test. So we do not go that far into it and the kids really do want to know more about certain people and certain dates in history but it's not on the test. [The students] don't show their creativity (Teacher H-1, from high SPS/ higher SES school).

It's less fun - you have to bring it in. We are doing some storyboards - the kids are loving every second of it and I'm thinking we really probably should have waited till after the test and that is not right. They are learning things. They are going to have to enjoy learning after the test (Teacher H-2, from high SPS/ higher SES school).

This finding of non-tested subjects that are neglected or postponed is consistent with the findings of Smith et al. (1989). They reported that high stakes testing caused some teachers to "neglect material that the external test does not include...reading real books, writing in authentic context, solving higher-order problems, creative and divergent thinking projects, longer-term integrative unit projects, [and] computer education..." (p. 268).

2. How much time do teachers perceive that students spend on test preparation and how does that amount of time compare to the time spent on instruction?

### Time

Besides the neglecting of non-tested subjects, the teachers interviewed stated that one of the most glaring impacts of LEAP/iLEAP testing was on instructional time. To comprehend this effect one should understand two other conditions of a normal classroom environment: a crowded curriculum (the volume of content area mandated by the state, parish, and school

require teachers to address is beyond the capacities of teachers to cover comprehensively), and the fixed number of teaching hours available. Time is a restricted resource and is systematically limited when LEAP/iLEAP testing is considered.

Under these circumstances elementary teachers have had even greater time restraints because generally they must teach every subject area, and the elementary curriculum usually includes lessons in areas in addition to academic ones (e.g., social skills). According to the participants of the study this outcome resulted in educators rushing to try to transform several different objectives into daily lessons, and then fit everything else into the allotted time.

Teachers' planning time is also diminished by numerous staff meetings concerning testing strategies; administrator discussions with teachers on ways to improve test scores; group and individual trainings regarding testing; and instructing teachers with materials to improve their students' test-taking skills.

[Due to the amount of time available] the way I do it is I plan my lessons out [to cover everything] and I feel that it is important that for the children - it might be hit or miss - but at least they were exposed to it (Teacher C-2, from high SPS/lower SES school).

The test taking material touched on topics that we wouldn't get to till May. We may not logically get into fractions till May but to do 8 lessons of learning by the time LEAP comes around and the test prep materials - we had to speed it up so that at least they have seen it. Speed it up and cram it so they may have seen it. It's drugged through the mud; some are going to do well with it. The others [students] that are not ready for it might not get it all the way (Teacher D-1, from high SPS/lower SES school).

I feel like I am on a time clock (Teacher E-2, from poor SPS/ higher SES school).

[W]ith the test prep we've not had enough time to review addition and subtraction [but] we set aside time everyday - we go ahead and do that (Teacher F-2, from poor SPS/ higher SES school).

I think that the Parish keeps grabbing programs to improve test scores and we have so many new programs and we test so many times between LEAP and interval assessments that are supposed to improve LEAP that we as teachers don't have time to teach. We test so much that it takes away from teaching time. We don't get to teach (Teacher G-1, from high SPS/ higher SES school).

I can't go back and re-teach because of the time. I have to move on even if ¼ or half of my kids don't get it. I have to move on. In science there are a lot of topics and more hands-on experiments but we don't have time, this year I've only done 2 or 3. It is to the detriment of the children (Teacher G-2, from high SPS/ higher SES school).

... I do have to speed it up (Teacher H-1, from high SPS/ higher SES school).

These findings confirm previous results that concern the degree of instructional time teachers devote to test preparation. Smith & Rottenberg (1991) reported on average, teachers spent **3** hours of test preparation for every hour of high stakes testing administration. Smith et al. (1989) reported that teachers in two case study schools spent up to four weeks of school time on test preparation for high stakes tests, and that time spent on test preparation increased as the test date became closer.

### Fairness

A concern of the teachers that was revealed during the interviews concerned the fairness of the test. Some of the teachers were concerned about the utility of the LEAP tests and its suitability for some students. These findings raise possible questions concerning inconsistencies between teachers' perceptions about the utility of LEAP/iLEAP testing and their classroom instructional behavior concerning test preparation that will be addressed later in this chapter.

Many of the teachers, especially those with past testing success, believed that their students had a clear disadvantage. This disadvantage was the result of being held to the same standard as wealthier children who have more resources and advantages.

You know our economic area [poor] it's a negative when it falls short you have all the higher ups coming to your school with no explanations (Teacher C-2, from high SPS/lower SES school).

I'm a Special Ed teacher. My children are forced to meet the same demands [as the wealthier schools] ... its just frustration I'm forced to teach them something and hold them accountable for the test [when their situation is not the same] (Teacher D-2, from high SPS/lower SES school).

I feel bad for Special Ed and disabled students who can't because it's all a higher level (Teacher G-2, from high SPS/ higher SES school).

Some the teachers thought it unfair that students coming from other situations were counted against the school's SPS score. The consternation resulted because they felt they should not be held accountable for students they did not have time to teach.

I think the teachers work very hard. It's not like the students do bad because the teachers aren't teaching. Everyone has their nose to grindstone. If you walk around you will see it for yourself. Where the problem lies, I think, the kids come to us from where they start. We have such a long haul [to catch them up] (Teacher A-1, from poor SPS/lower SES school).

It makes Louisiana look so ignorant because when you get students who transfer in - their kids are never taught and their tests are different from our tests. They are ready for their tests but not ready for our tests... or you get the kids from the Catholic schools who aren't teaching the tests. So, they don't really care and they are good on phonics and decoding but you ask them to write because our tests are focused on that and they can't write because when they come to you they are lost (Teacher G-2, from high SPS/ higher SES school).

Some of the teachers were frustrated with how the SPS scores were calculated. The fact that the fourth grade was held to the same standard as the fourth grade the year before, ignoring the differing capabilities of the two classes.

You're comparing apples to oranges ... you have classes that are bright and then classes that are not as bright. Comparing my 4<sup>th</sup> graders of last year to my 4<sup>th</sup> graders of this year - that's not fair because you're not following them to watch them grow (Teacher G-1, from high SPS/ higher SES school).

[When you compare last year's class to] this year's you are comparing apples and oranges. I don't know what they compare from year to year. I think it should be progressively looking at the same class and maybe compare from third to fourth (Teacher H-1, from high SPS/ higher SES school).

[You wonder] can you keep going higher are you going to be able to match last year's scores and it's different kids (Teacher H-2, from high SPS/ higher SES school).

Something that emerged from the interviews was the practice of moving students to the next grade level that are not ready. This is a side effect of the LEAP test.



A child can only repeat once between Kindergarten and 3<sup>rd</sup> grade, but if they need to be retained again - don't worry because when they get to 4<sup>th</sup> grade they will be kept back. Instead of keeping them back in 2<sup>nd</sup> grade with the basics. Wait to catch them in 4<sup>th</sup> grade. By the time they get to 4<sup>th</sup> grade they are not going to write 150 words if they can't write 30 words. I have a problem with children being forced to move up and catch it in 4<sup>th</sup> grade (Teacher E-2, from poor SPS/ higher SES school).

### Focus on Instruction

Not all the responses to the interviews concerning instruction were negative. Some teachers felt that testing did force educators to focus on students who had been ignored. Also, some teachers thought testing was needed to focus uncommitted teachers, although they did not put themselves in that category. However, one teacher believed that this focus was not producing more knowledgeable students, just better test takers.

Some of the weaker teachers have been asked to give up that grade [4<sup>th</sup> grade] and some have no other choice because if you are at the bottom of the totem pole that's it. [The teacher was referring to teachers whose students were doing poorly on the LEAP test being fired. She presented this as positive result] (Teacher B-2, from poor SPS/lower SES school).

I like the philosophy of it [LEAP testing] many kids are addressed now that weren't before. They were called the 1013 kids because they lived next door. I think that school just dismissed them and now they can't dismiss them (Teacher C-1, from high SPS/lower SES school).

I'm just handing out stuff – it's an assembly line. I just I feel I had no control over what I thought was beneficial for the kids which I understand can be a good thing because I think my judgment is perfect but other people may not have that perfect judgment (Teacher C-2, from high SPS/lower SES school).

I think LEAP has been a good thing. For one thing it has made the parents more involved like if you send home a packet in fourth grade, the parents make sure it is done (Teacher G-1, from high SPS/ higher SES school).

and

We should be held accountable for [creating] better well rounded students but all we are doing is providing test kids. I think it's headed in the right direction but I think the goal that they are trying to accomplish is getting them [students] to learn only the basics (Teacher D-2, from high SPS/lower SES school).

## Pressure

The advent of LEAP/iLEAP testing saw an increased amount of pressure placed on teachers. This reality was confirmed by both the surveys and the interviews conducted in this study. It was also documented in research reported in Chapter 2. The pressure on teachers proved to be from many sources and at many levels. At one level teachers experienced not only their own pressure but were empathetic to the students' pressure. Some teachers also thought this affected the way they saw the students; because they started to look at their students as high scorers or low scorers.

At another level, the teachers experienced their own pressure from the State, Parish, administrators, parents, and the media. Astonishingly, the surveys revealed that teachers in schools with a history of high SPS scores reported feeling more pressure from every source than teachers in schools with a history of low SPS scores. This data was surprising because of the prevailing myth among many in education that if you could just make your target scores that all stress and pressure would go away. However, the opposite was found to be true. As the target SPS scores escalate, the amount of pressure to push students and teachers beyond their capabilities also soars. The interviews helped to illuminate this phenomenon.

### Students' Pressure

Many teachers perceived that LEAP testing was malicious and cruel because of the pressure that it placed on the young students. The length and difficulty of tests, the number of tests (including practice tests), and the high stakes nature of the test led teachers to believe that LEAP/iLEAP caused stress. Some teachers believed that the tests caused their students to develop test anxiety.

We had more to do. It put more pressure on us to make sure they were ready for the tests and for Special Ed because they were behind or had exceptionalities. They felt the pressure of how far [they are behind]. They got more frustrated to obtain the knowledge for the LEAP tests. The teachers felt pressure too. I would eliminate the pressure from

how to teach it to them by teaching on their level instead of using the materials by the State. I tried to teach the way they wanted them to learn and they couldn't understand it and it made them frustrated because they couldn't understand it. It became harder to individualize the lessons because they are on different levels and the entire class is focused on tests ... It shows what kids can do but it brings them down. It has been a detriment. It causes a lot of kids to drop out like a senior who passes all of their classes but misses the test by 3 points and can't cross the stage, in elementary school I know 2 kids who didn't want to go to school because of LEAP and they are in jail now (Teacher A-2, from poor SPS/lower SES school).

I think the pressure on the teachers trickles down to the students (Teacher B-2, from poor SPS/lower SES school).

[The thing that is most negative] I think is the stress on the kids. I had to break the news to a fourth grader. He burst into tears because he didn't pass. Honestly, it was heart breaking (Teacher C-1, from high SPS/lower SES school).

I think there is way too much testing of the kids: weekly tests, test interventions, first interval assessment, regular test prep, and LEAP testing for the kids. It is stressing them out (Teacher F-1, from poor SPS/ higher SES school).

I think they test the kids a little too much. It's a lot of pressure for them (Teacher F-2, from poor SPS/ higher SES school).

I think its way too much pressure for a child in the fourth grade. To pass a test to go onto the next grade, I think, you know, a ten year old student should not be put under that pressure. I think the discipline in fourth grade has gone down. Therefore, the kids feel the pressure and the teachers are stressed to get them (Teacher D-2, from high SPS/lower SES school).

I just think it's too much pressure on a 9 year old ... it's too much pressure on those kids. I'd rather the pressure be on me rather than them. It was stressful worrying about my own two 4<sup>th</sup> graders. You try to relax them but they know if they don't pass it is back to the fourth grade. How well you do kind of follows you ... (Teacher E-1, from poor SPS/ higher SES school).

I just think for children who don't pass it is horrible on their self esteem. It dictates their future (Teacher G-1, from high SPS/ higher SES school).

### Teachers' Pressure

When I interviewed teachers about the pressure that they had experienced many of the interviewees were very emotional. These teachers felt ashamed and embarrassed if their students scored low or failed to meet district standards. Some experienced relief rather than pride when scores were elevated. This internalized pressure sometimes manifested itself in their relationships

with their colleagues.

I think the State put the pressure because they were here. It came from the State and central office because they were always there and watching ... It kind of made us feel like what else could we do? What else can I do to help? It made us feel like we weren't doing enough (Teacher A-2, from poor SPS/lower SES school).

We didn't do well so more pressure. It's been an issue since I've been here but the worse the scores get the more it gets. I'm sure central office is throwing it to the principal but wherever it is coming from it is a lot. It is coming from the principal (Teacher B-1, from poor SPS/lower SES school).

Before testing the kids are ready to crack and the teachers are ready to crack (Teacher G-2, from high SPS/ higher SES school).

Besides the State, Parish, and administrators ... the media I do feel puts a lot of pressure on us by publishing the names and school and scores and you know they have every right to. They also write the article against education. We get a lot of it from bosses and from the media. This is a basic skills test. And if we just taught the basics then we would be fine. It is added pressure. My first time teaching fourth grade, only 7 of them [her students] were scored in the "Unsatisfactory" [category] for Language Arts. I felt 7 out of this many, yeah but that's 7 you obviously didn't teach well. But that's 7! (Teacher H-2, from high SPS/ higher SES school)

This loss of autonomy was also a contributing factor to a stressful environment.

I would say it was a negative thing because we had so much coming at us and it was like put the focus here, no here and it could easily become frustrated and overwhelming. I got 18 things to do and guess what? I can't do any of it (Teacher E-2, from poor SPS/ higher SES school).

To me school is not as pleasant as it used to be. I mean it's just not as pleasant as it used to be. And I don't think that's a good thing. You know school should be fun. It shouldn't be from the time you walk in - you know - y'all got to learn this data, data, data but that's what happens. I will say this, there is some tension among teachers because the 4<sup>th</sup> grade teacher is responsible for the LEAP test and they go and blame the third grade teacher. Not animosity [among teachers] but there is friction (Teacher C-1, from high SPS/lower SES school).

An interesting phenomenon discovered in the quantitative portion of this study was that schools that have a history of scoring high on the LEAP test are not absolved of pressure. To the contrary, the amount of pressure increased due to higher expectations. High scores induce the SPS target score to grow exponentially for the next year. This produces a greater amount of stress and pressure on the teachers. Take note that one of the teachers interviewed actually felt

such an amount pressure that she honestly rooted for students to fail in order to lower expectations and the target score. Of course, as stated earlier, failing the LEAP test does result in the child repeating 4<sup>th</sup> grade.

I think sometimes its demoralizing in a way, only because there's too much pressure on you. We get kids all the time from other schools and they're not prepared. How could you let this happen? But we are held responsible for the student. Yeah, this is specific to our school [last year] was a baseline year. **I hoped everyone didn't pass.** You don't want to set the bar to high but you want to leave room for improvement. People start complaining (Teacher C-1, from high SPS/lower SES school).

The SPS - doesn't it have to keep increasing? Of course there's always pressure to better yourself. We don't want to get in any corrective action. Even if you're doing well you can be put into corrective action... It is too much (Teacher D-2, from high SPS/lower SES school).

There is added pressure because we are always expected to be in the top. I think what is really unrealistic is that if we were the number one school for a year with a score of 120 and then we score a 115; we become a school in corrective action. This means that they are going to come in and teach us how to teach. But if you were a school of 50 that moves to 55 they leave you alone. That doesn't make sense ... We weren't number one this year and I felt very bad, we were up there but not number one this year. I don't know if we felt down but we weren't number one. There was more pressure to make sure we don't keep falling - maybe just pressure with ourselves (Teacher G-1, from high SPS/higher SES school).

As evidenced in the examples above, the teachers are feeling pressured and besieged because of the great emphasis that is placed on LEAP/iLEAP testing and the pressure they feel to raise the scores. This pressure comes from many sources and in many forms. In fact much of the pressure was self – imposed from the teachers.

3. What effect does testing have on an educators' sense of professionalism and pride in their work? How does high stakes testing affect motivation in general?

#### Commitment

The third hypothesis, stated in Chapter 4, was that teachers' degree of commitment would depend on the school's SPS scores (poor vs. high). An ANOVA indicated a significant difference did not exist in the degree of commitment between schools scoring high and schools scoring

poor on the SPS. The variable of commitment seemed to be a very individualized and personal one that was not overtly affected by past testing success or socioeconomic conditions. With the qualitative data we will discover other possible factors that affect commitment.

As stated earlier, the teachers are forced to use the GLEs as a pacing guide. However, they regard this pace to be completely unrealistic. Furthermore, they were puzzled that the LEAP was based on the traditional calendar year yet the test was administered in March. The GLEs forced teachers to compress into six months what they would usually teach over a nine-month period.

This fact, along with many other factors discussed previously, led to stress and increased pressure on the teachers. However, for some teachers it went further than increased stress. This loss of autonomy also affected their commitment to the educational profession.

An unintended consequence is that teachers are leaving public education to take positions in private schools to avoid the pressure of the high stakes tests. In order to reduce the constant stress of the pressure to spur high student achievement on a single measure of knowledge, highly qualified teachers are leaving to teach in private schools which are free of state mandated tests that make them feel compromised as professionals (Amrein & Berliner, 2003). These unintended results have impacted every aspect of the American education system (Buck, 2006, p. 4).

Many teachers took exception to the instructional mandates because they were no longer allowed to use their professional judgment and experience to design individualized instruction.

I'm not really losing control but a lessening of control of what you can and cannot do. Practically all of it is directed to the LEAP tests (Teacher A-1, from poor SPS/lower SES school).

Test prep in a normal class 50% to tests and 50% to overall. It goes together. 50% is test taking skills. I really don't have much say, so I'm told what to teach and that's what I taught (Teacher B-1, from poor SPS/lower SES school).

... I felt I had no control over what I thought was beneficial for the kids (Teacher C-1, from high SPS/lower SES school).

I'm a professional 4<sup>th</sup> grade teacher, if it weren't for LEAP I would teach in a different sequence ... I do get fearful that some of them just aren't going to be able to get it by then (Teacher D-1, from high SPS/lower SES school).

UNO [outside consultants] gave the first graders 100 problems - addition and subtraction. Most of the kids got low scores. The test isn't a good test to give to them ... I don't think that they are very good at interventions either. Flashcards are a better intervention for them than writing on the board but they tell us that's what we have to do so (Teacher F-1, from poor SPS/ higher SES school).

In a way I do lose control what I can teach. Our principal is very much into pushing the test, but she doesn't come in and see. I feel I have a little leeway because she isn't on top of me. So, I can kind of get away with a little creative stuff but if there was no test it would be different. I wouldn't be teaching what I'm teaching (Teacher G-2, from high SPS/ higher SES school).

... [O]ne day they're going to give us a script; this is what you say; this is what you ask; and move on (Teacher H-2, from high SPS/ higher SES school).

I think it is dumbing down for the teachers, I think this is showing less respect for the teachers because it's saying your professional judgment is not adequate. It has taken some of the choice from the teachers (Teacher C-1, from high SPS/lower SES school).

Many of the teachers believe that they are being disregarded. They are forced to apply one-size-fits-all strategies and curriculum. They are following guidelines and pacing charts that they believe are inappropriate for their students. They are frustrated with being compared with teachers of higher socioeconomic communities and being blamed for their students' low scores when other contributing factors exist. This belief is supported by Berk (1988) who discovered that high stakes tests correlate considerably with many student characteristics other than program or teacher quality.

High-stakes tests may motivate certain teachers and some students to achieve optimal performance levels. However, researchers have cautioned that placing a premium on student test performance can lead to instruction that is reduced primarily to test preparation, thus limiting the range of educational experiences for students and constraining the pedagogical skills of teachers. Studies have also shown that high stakes assessments increase stress and decrease morale among teachers. More than 77 percent of North Carolina teachers surveyed indicated decreases in their morale; 76 percent reported that teaching was more stressful since the implementation of the North Carolina state testing program. In Texas, 85 percent of teachers surveyed agreed with the statement 'some of the best teachers are leaving the field because of the TAAS [Texas' version of the LEAP test]' (Abrams, 2004, p. 8).

The teachers from schools with a history of high SPS scores are also irritated with the fact that they feel they are being punished for prior success with the elevation of target scores. These factors are contributing to a lessening of commitment from some educators, especially the younger ones who have less of a stake in the profession and retirement benefits.

I'm a Special Education teacher my children are forced to meet the same demands [as other students] ... absolutely, meaning that I don't know if I'd want to continue to teach. I like the education field. I like being in Special Ed. It's just frustration. I'm forced to teach them something ... even if you're doing well you can be put into corrective action. My feelings are to leave the teaching profession (Teacher D-2, from high SPS/lower SES school).

[Due to being pressured] I was ready to quit at my other school (Teacher E-1, from poor SPS/ higher SES school).

Just this week, I was talking to one of these older teachers here that I don't know if I can keep doing this for 20 or 30 years (Teacher F-1, from poor SPS/ higher SES school).

I find that the parish takes on so many programs. I don't think that we're less committed but I think we may be more likely to tell someone to not go into the profession (Teacher G-1, from high SPS/ higher SES school).

The frustration hasn't affected my commitment or motivation as a teacher, but I know two teachers on sabbatical who finally had enough and I don't know if they are going to come back due to the tests. My daughter is in education getting ready to graduate in May and I keep saying it's not what I thought it was going to be like. It's getting more every year. They are throwing in more paper. I mean with the behavior you're tracking. I feel like I have control over my classroom, but then I have to do all this stuff. They are telling you to give this test, this pretest, and this posttest. I feel you don't teach and there is more paperwork (Teacher G-2, from high SPS/ higher SES school).

These quotations are consistent with the findings of Johnson & Johnson (2006). In their observations they noted many of the newer teachers were leaving the educational profession due to LEAP testing and that the older teachers were counseling interested parties to choose other professions. The answers to the Research Questions are summarized in Table 5.3.

These quotations listed above would seem to support Self-Determination Theory's perspective on testing. This theory was discussed in Chapter 2.



## Summary of Chapter 5

This mixed methodology study was designed to provide data regarding the perceptions of Jefferson Parish teachers toward LEAP 21 Testing and how high stakes testing affected school improvement. In Phase II a survey (Chapter 4) provided baseline data on the study sample. Phase III was informed by the results of Phase II. Teachers were chosen from those teachers that participated in the surveys. During Phase III, the researcher utilized Multistage Cluster sampling. This sampling technique allowed the sample to be further reduced within the eight clusters (schools) to identify teachers who could add knowledge to the interview. Normally, the Multistage Cluster sample is random; however, it added richness to the data for the researcher to identify teachers through their survey responses who provided the most informative answers to the research questions. Two teachers were interviewed from each school for a total of sixteen teachers. On-site visitations were conducted for the purpose of gathering qualitative data through interviews.

The subjects were discussed and the Interview Guide Approach was employed to add depth and understanding to this study. The interviews were conducted and the constant comparative method of analysis was utilized. Patterns and themes of perceptions of participants emerged lending data that triangulated with the survey data. Themes involving instruction, pressure, and commitment were evident and discussed.

The following four points summarize the results of the qualitative analysis:

- (1) The qualitative responses helped explain the results from the quantitative results. They were complementary in nature.
- (2) Most teachers really dislike the high stakes testing regardless of how well their students do on the test and or what their (the students) SES level was.

Table 5.2:

Hypotheses Results

<b>Prediction</b>	<b>Quantitative Results</b>	<b>Qualitative Results</b>
<p>Jefferson Parish teachers from schools that produce high SPS scores will perceive that LEAP 21 testing has affected classroom practice more than teachers from schools that have scored poorly.</p>	<p>The surveys failed to confirm the hypothesis. The ANOVA table indicates a significant difference in perceived effect on classroom practices between schools scoring high and schools scoring poor on the SPS does not exist. From Table 4.3 of this study, we see that schools scoring higher indicated a higher effect on classroom practices. Note that in this table, although all of the categories scored high, Poor SPS/ Lower SES cell reported somewhat lower on classroom practices; whereas, the rest of the categories reported similarly high scores. This one cell accounts for the interaction.</p>	<p>Reaffirmed quantitative results. There was little difference in the four cells, using SPS and SES as the independent variables, created for this study. The themes of teaching to the test, neglecting subjects, time, fairness, and focus on instruction were discovered (see p. 109).</p>
<p>Jefferson Parish teachers from schools which produce low SPS scores will perceive that LEAP 21 testing has created pressure to spend more time on test preparation (teaching to the test) than teachers from schools that have achieved high SPS scores.</p>	<p>The surveys failed to confirm the hypothesis. The ANOVA table indicates no significant difference in perceived pressure between schools scoring high and schools scoring poor on the SPS. From Table 4.5 of this study, we see that schools scoring higher indicated a higher perceived pressure to spend more time on test preparation.</p>	<p>Reaffirmed quantitative results. There was little difference in the four cells, using SPS and SES as the independent variables, created for this study. The themes of student and teacher pressure were discovered (see p. 118).</p>
<p>Teachers from schools with high SPS scores will indicate that they have a higher degree of commitment to the education profession than teachers from schools with low SPS scores.</p>	<p>The surveys found this hypothesis to be false. The ANOVA table indicated no significant difference in degree of commitment between schools scoring high and schools scoring poor on the SPS.</p>	<p>Other factors were introduced and some teachers were more affected by testing than other teachers. The theme of commitment involving the issues of loss of control and professionalism was discovered (see p. 121).</p>

Table 5.3:

Research Questions Results

<b>Research Questions</b>	<b>Quantitative Results</b>	<b>Qualitative Results</b>
<p>How does test preparation (teaching the test) affect teachers' instructional planning, learning strategies, and curriculum content and to what extent?</p>	<p>The surveys did not address this question.</p>	<p>All of the teachers interviewed stated that LEAP 21 testing did affect their instructional planning, learning strategies, and curriculum content. To what extent they were affected varied but it was even more prevalent in schools with a history of high SPS scores. Such practices as teaching to the test, neglecting subjects, sequencing, and time allotment were discussed (see p. 109).</p>
<p>How much time do teachers perceive that students spend on test preparation and how does that amount of time compare to the time spent on instruction?</p>	<p>The surveys did not address this question.</p>	<p>All of the teachers interviewed stated that LEAP 21 testing forced them to devote some time to test preparation. This was time that would normally be used for instruction. The amount of time devoted to test preparation grew exponentially as testing neared. Teachers provided a range of 1/3 of class time to a 100 %.( see p. 113).</p>
<p>What effect does testing have on an educators' sense of professionalism and pride in their work? How does high stakes testing affect motivation in general?</p>	<p>The surveys did not address these questions specifically.</p>	<p>Many of the teachers felt internalized pressure that sometimes manifested itself in their relationships with their colleagues. The loss of autonomy was also a contributing factor to a stressful environment. . Many teachers took exception to the instructional mandates because they were no longer allowed to use their professional judgment and experience to design individualized instruction. Furthermore, teachers were frustrated with being compared with teachers of higher socioeconomic communities and being blamed for their students' low scores when other contributing factors existed. These factors are contributing to a lessening of commitment to the educational profession from some educators, especially the younger ones who have a less of a vested interest in the profession. (see p. 121).</p>

- (3) There were several themes that ran through the teachers' responses including those summarized in Tables 5.2 and 5.3. These themes were consistently voiced by teachers from all the categories.
- (4) The study started out as a QUAN+qual mixed design, but wound up as a QUAN+QUAL or perhaps evens a quan+QUAL design. This is an example of how it is difficult to determine the priority of methodological type before the study is before it is conducted (e.g., Teddlie & Tashakkori, 2008).

The discussion and recommendations that are informed by Chapters 4 and 5 will be presented in the next chapter.

## CHAPTER 6: SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

### Introduction

This study that utilized both quantitative and qualitative data collection techniques was devised to provide information regarding the opinions of Jefferson Parish teachers toward LEAP 21 testing and how high stakes testing impacted school improvement. Explicitly, I researched the experiences and perceptions of the teachers regarding the effects of high stakes testing in the areas of classroom practices, pressure caused by the inception of testing, and commitment to future employment in the Jefferson Parish Public School System. Teachers' experiences and perceptions were explored through surveys and interviews. Chapter 6 commences with the summary of the study and discusses the common patterns and themes that emerged during the analysis of the data. The data analysis is followed by implications of the study and recommendations for future research.

### Summary of the Study

This study was organized as an investigation of the effects of high stakes testing on the area of commitment, pressure, and classroom practices as perceived through the experiences of teachers. The research design that was utilized for this study was a mixed methods design that was divided into three phases to verify and explore high stakes testing's effects on school improvement and daily operations.

Phase I utilized previous surveys and a peer review to create a knowledge base to generate a survey instrument that was focused on the three areas of this study. Simultaneously, elementary schools were identified as having past success or failure concerning LEAP 21 testing by utilizing data provided by the Louisiana Department of Education. Finally, the survey instrument that was created was piloted in three Jefferson Parish middle schools to identify strengths and weaknesses.

Phase II comprised the dissemination of surveys to all the teachers of eight Jefferson Parish elementary schools identified through sampling. These data were divided into three dependent variable groups, each statistically analyzed to determine if significant differences existed between schools across two independent variables: SES and past testing success.

Phase III was informed by the results of Phase II. Teachers were chosen from those teachers who completed the surveys. During Phase III, the researcher utilized Multistage Cluster sampling. This sampling technique allowed the sample to be further reduced within the eight clusters (schools) to identify teachers who could add knowledge to the interview. Normally, the Multistage Cluster sample is random. In this case; however, the researcher identified teachers who had provided the most informative answers to the survey questions, thereby generating a sample that would generate rich data. Two teachers were interviewed from each school for a total of sixteen teachers. On-site visitations were conducted for the purpose of gathering qualitative data through interviews.

The quantitative section (Chapter 4) attempted to address the hypotheses put forth in this study. The first hypothesis was that teachers' perception of the LEAP testing effect on classroom practices would depend on the school's SPS scores (poor vs. high). A Two-Way ANOVA determined that an interaction between the two independent variables existed meaning that the difference in response between SPS high and SPS poor depends upon the socio-economic level of the school. This was also evident through the means and standard deviations collected for this report. Although all of the categories scored high, Poor SPS/ Lower SES cell reported somewhat lower on classroom practices; whereas, the rest of the categories reported similarly high scores. This one cell accounts for the interaction.

The second hypothesis regarded teachers' perception of whether LEAP testing created pressure to spend more time on test preparation would depend on the school's past SPS scores

(poor vs. high). Surprisingly, Table 4.5 demonstrated that schools scoring higher indicated a higher perceived pressure to spend more time on test preparation than the poor scoring schools.

The third hypothesis was that teachers' degree of commitment would depend on the school's SPS scores (poor vs. high). An ANOVA indicated no significant difference in the degree of commitment between schools scoring high and schools scoring poor on the SPS. The means and standard deviations collected for this report indicated that teacher's placement in a school scoring poor or high on the LEAP 21 test or high/low SES had little effect on their commitment to the profession. The effects were more individualized. To better understand the impact of LEAP/iLEAP testing, a qualitative analysis was introduced.

The results of the qualitative section attempted to address the research questions that were presented in this study. The first question inquired about the extent that teachers' instructional planning, learning strategies, and curriculum content were affected by test preparation (teaching the test). The teachers reported that the Parish and State's emphasis on the LEAP/iLEAP had resulted in drastic changes to the curriculum. A majority of the teachers stated that in creating their lesson plans they look at prior LEAP tests to make sure that their curricula includes all or most of the test content. They also reported a greater emphasis on the content that would be tested and subjects not tested were often ignored. Further, they report that the test forced them to focus on breadth more than on depth of coverage of the material. The majority of teachers also adjust the sequence of their curriculum based on what is included in the test. Some teachers didn't change the sequence. They, however, were forced to cover some subjects in a less comprehensive manner. In addition to these changes in what was taught, most interviewees felt that preparation for the LEAP/iLEAP test had produced changes in how they taught.

The second question involved the amount of time that teachers and students spent on test preparation compared to the time spent on instruction. Besides the neglecting of non-tested

subjects, the teachers interviewed stated that one of the most glaring impacts of LEAP/iLEAP testing was on instructional time. The teachers considered class time a restricted resource that was systematically limited when LEAP/iLEAP testing was introduced. The teachers' planning time was also diminished by numerous staff meetings concerning testing strategies; administrator discussions with teachers on ways to improve test scores; group and individual trainings regarding testing; and instructing teachers with materials to improve their students' test-taking skills.

Other subjects that were raised by the interview process were that some the teachers thought it unfair that students coming from other situations were counted against the school's SPS score. The consternation exists because they felt they should not be held accountable for students they did not have time to teach. Some of the teachers were frustrated with how the SPS scores were calculated. The fact that the fourth grade was held to the same standard as the fourth grade the year before ignored the different capabilities of the two classes. Furthermore, something that emerged from the interviews was the practice of moving students to the next grade level that were not ready. This was done because the administrators had projected the students to repeat fourth grade anyway.

Not all the responses to the interviews concerning instruction were negative. Some teachers felt that testing did force educators to focus on students who had been ignored. Also, some teachers thought testing was needed to focus uncommitted teachers, although they did not put themselves in that category. However, one teacher believed that this focus was not producing more knowledgeable students, just better test takers.

The third question inquired about the effect that testing has on an educators' sense of professionalism and commitment to continuing in their profession. The surveys and interviews both demonstrated an increased amount of pressure on teachers to raise scores in all situations.



The pressure on teachers proved to be from many sources and at many levels. At one level teachers experienced not only their own pressure but were empathetic to the students' pressure. Some teachers also thought this affected the way they saw the students, because they began to look at their students as high scorers or low scorers.

At another level, the teachers experienced their own pressure from the State, Parish, administrators, parents, and media. Astonishingly, the surveys revealed that teachers in schools with a history of high SPS scores report feeling more pressure from every source than teachers in schools with a history of low SPS scores.

Many of the teachers felt ashamed and embarrassed if their students scored low or failed to meet district standards. Some experienced relief rather than pride when scores were elevated. This internalized pressure sometimes manifested itself in their relationships with their colleagues. The loss of autonomy was also a contributing factor to a stressful environment.

An interesting phenomenon discovered in the quantitative portion and illuminated in the qualitative portion of this study was that schools that have a history of scoring high on the LEAP test are not absolved of pressure. To the contrary, the amount of pressure increased due to higher expectations to the point that one teacher admitted to rooting against all of her students passing.

Teachers also regard the pace set by testing to be completely unrealistic. Furthermore, the GLEs forced teachers to compress into six months what they would usually teach over a nine-month period. Many teachers took exception to the instructional mandates because they were no longer allowed to use their professional judgment and experience to design individualized instruction.

Furthermore, teachers were frustrated with being compared with teachers of higher socioeconomic communities and being blamed for their students' low scores when other contributing factors existed. These factors are contributing to a lessening of commitment to the

educational profession from some educators, especially the younger ones who have less of a vested interest in the profession.

Kate Randall (2001) states:

Many school districts—especially those in low-income, urban areas—are already hard-pressed to recruit and train qualified teachers. In New York City, the Board of Education still needs to hire 3,300 teachers before schools reopen in September. The board has conducted an \$8 million ad campaign this summer to recruit new teachers, and only 56 percent of these 4,700 new teachers are certified. More than 7,000 teachers are expected to leave the city’s schools next year....Difficult working conditions for teachers in many school districts—including severely overcrowded classrooms and a lack of resources—will only be exacerbated by pressure on teachers to achieve passing scores from their students on the annual standardized tests (p. 2).

#### Discussion

LEAP testing has been in place since 1999. Everyone interviewed agreed that it has had an impact like no other mandate from the State. Schools have been disbanded; teachers and principals reassigned or fired; and students who have passed their courses have been forced to repeat the grade due to a single, standardized test. The results from these tests have been given a massive amount of influence in the educational system by policy makers. This influence has been granted despite the inadequate amount of research concerning the LEAP 21 Test, especially studies involving those closest to witness the daily affects of this policy.

Despite this lack of research, the influence that is given to this form of assessment has not weakened over time. The opposite is true. Recently, in December, the Orleans Recovery District Superintendent, Paul Vallas, proclaimed a plan to assess bonuses of up to \$3,000 to teachers and up to \$5,000 to principals if low-performing schools in the state-run district achieve a specific school performance score (SPS). By implementing the program this year the superintendent hopes to “... improve chances of getting additional federal money earmarked for recruitment and retention in hurricane-affected regions or tapping a separate source of federal dollars for performance-based programs” (Simon, 2008, p. B-1).

Subsequently, the policy makers are putting their faith and support behind this form of

high stakes testing. However, this study has reinforced previous studies (Amrein & Berliner, 2002; Kohn, 2000; Pedulla et al., 2003; and Wright, 2002) that have produced results which show side effects to high stakes testing that teachers believe are detrimental to their students.

The results of this study did indicate that the teachers perceived that one of the goals of LEAP testing (i.e., providing a focus on students who were “Left Behind”) was taking place. However, the teachers in this study felt that this accomplishment came at a cost. Most teachers increased the instructional time dedicated to tested subjects at the expense of non-tested subjects and extracurricular activities. In addition, a great source of apprehension among teachers was that a considerable share of instructional time was devoted to test preparation. Teachers also expressed concerns regarding the fairness of the test.

Not only are teachers’ perceptions concerning LEAP testing’s negative consequence on class instruction disturbing, the perceived impact of the testing on the teachers is also troublesome. The results of this research suggest that both students and teachers experienced increased pressure from many sources.

An interesting trend discovered in the quantitative portion of this research and supported by the qualitative portion was that schools that have a history of scoring high on the LEAP test experienced great pressure. The amount of pressure increased due to higher expectations. High scores induce the SPS target score to grow exponentially for the next year. This produces a greater amount of stress and pressure on the teachers. One of the teachers interviewed actually felt such an amount of increased pressure that she honestly rooted for students to fail in order to lower expectations and the target score. This occurred despite, as stated earlier, failing the LEAP test does result in the child repeating 4<sup>th</sup> grade.

I think sometimes its demoralizing in a way, only because there's too much pressure on you. We get kids all the time from other schools and they're not prepared. How could you let this happen? But we are held responsible for the student. Yeah, this is specific to our school [last year] was a baseline year. **I hoped everyone didn't pass.** You don't want to

set the bar to high but you want to leave room for improvement. People start complaining (Teacher C-1, from high SPS/lower SES school).

There is added pressure because we are always expected to be in the top. I think what is really unrealistic is that if we were the number one school for a year with a score of 120 and then we score a 115; we become a school in corrective action. This means that they are going to come in and teach us how to teach. But if you were a school of 50 that moves to 55 they leave you alone. That doesn't make sense ... We weren't number one this year and I felt very bad, we were up there but not number one this year. I don't know if we felt down but we weren't number one. There was more pressure to make sure we don't keep falling - maybe just pressure with ourselves (Teacher G-1, from high SPS/higher SES school).

As evidenced in the examples above, the teachers are feeling pressured and beleaguered because of the great emphasis that is placed on LEAP/iLEAP testing and the pressure they feel to raise the scores. This pressure comes from many sources and in many forms. In fact much of the pressure was self – imposed from the teachers.

This increased pressure often resulted in a tense environment with a loss of autonomy. In many circumstances, these policies led to decreased morale and enhanced frustration. The result for some teachers of this frustration, tense environment, and loss of control over classroom decisions was to seek other avenues of employment. This is happening at a time when the school district is experiencing a severe shortage of teachers. This study also revealed that scoring well as a school on LEAP testing does not mitigate the pressures associated with testing because the expectations and target scores rise exponentially as a result.

The Louisiana District and School Accountability Advisory Commission (1998) representing the Louisiana State Board of Elementary and Secondary Education (BESE) established the following goals for LEAP 21 testing.

Louisiana's Public Education Accountability System is intended to drive fundamental change in classroom teaching by:

- Clearly establishing the state's goals for schools and students;
- Creating an easy way to communicate to schools and the public how well a school is performing;
- Recognizing schools for effectiveness in demonstrating growth in student achievement; and

- Focusing attention, energy, and resources on those schools that need help in improving student achievement (p. 1).

According to the majority of teachers surveyed and interviewed for this research, it is apparent that the anticipated goals of LEAP 21 testing policies set by policymakers contradict the aspirations of their implementation and have led to unintended negative consequences in the areas of classroom practices, pressure, and commitment to the educational profession. Therefore, it is paramount that policy makers reassess the high stakes nature of LEAP testing and the programs associated with it. Instead, officials have failed to acknowledge these unexpected negative consequences and, therefore, have done nothing to address them. Politicians and decision makers need to "...refocus education policies to place greater emphasis on supporting and improving teaching and learning, rather than relying on a system of rewards and sanctions to spur change in classrooms" (Abrams, 2003, p.27).

When the policy makers refocus their attention regarding the LEAP test, special attention should be given to teachers' concerns and input. Policy makers ignore the teachers at the detriment to the assessment program. As Jones and Egley (2004) state:

Until policymakers take teachers' concerns seriously and make an effort to address them, teachers will not likely support reform through high-stakes testing. Without the support of teachers, high-stakes testing will likely become just another failed education reform. However, with the input of those on the frontlines and some vital and well-conceived changes, testing programs are likely to have a more positive effect on the teaching and learning processes (p. 26).

By including teachers in the decision making, policymakers can revisit the original intent of this program. One of the original intents was to focus attention on low performing students. Ironically, they seem to be those most negatively affected by this program.

Studies have shown ... that such methods leave poorly performing students even worse off. For example, researchers at the Consortium on Chicago School Research have found that students who were held back had significantly higher dropout rates and lower achievement than those of similar ability who were promoted (Smith, 2001, p. 3).

The majority of the teachers interviewed for this study saw the LEAP 21 testing as less than perfect and had negative side effects for teachers, students, and schools, however they did indicate that if changes were made that it could be used in a positive way. The teachers believed that properly developed and used tests provide critical measures of students' progress and they did refocus attention on those students who are often ignored or forgotten.

I like the philosophy of it [LEAP testing] many kids are addressed now that weren't before. They were called the 1013 kids because they lived next door. I think that school just dismissed them and now they can't dismiss them (Teacher C-1, from high SPS/lower SES school).

I think LEAP has been a good thing. For one thing it has made the parents more involved like if you send home a packet in fourth grade, the parents make sure it is done (Teacher G-1, from high SPS/ higher SES school).

When the kids take the test, you do get to see if they are ready to move and what they can do and you can help them (Teacher F-2, from poor SPS/ higher SES school).

“In fact, the testing standards that APA helped develop (available at [www.apa.org/pubinfo/testing.html](http://www.apa.org/pubinfo/testing.html)) explain that appropriate testing can be key in identifying lower-performing students and schools so that they can get the extra resources they need” (Smith, 2001, p. 4).

Some teachers indicated that they would provide more support to LEAP 21 testing if it was used appropriately. Professional development would be key to promote understanding.

When tests are developed and used appropriately, they are among the most sound and objective knowledge and performance measures available. But, appropriate development and use are critical. Fairness in testing begins when tests are being developed. Test developers should provide to those using their tests (school systems, for example) specific information about the potential limitations of the test, including situations in which the use of the test scores would be inappropriate. For example, a test that has been validated only for diagnosing strengths and weaknesses of individual students should not be used to evaluate the educational quality of a school. Furthermore, those using a particular test should have an appreciation for how the test performance of some students--students with a disability or those with limited English-speaking ability, for example, should be interpreted (American Psychological Association, 2001).

Many proponents believe that if testing was presented with the appropriate support, professional development, and resources that it could contribute profoundly to school improvement (Elmore, 2000). Such improvements as making better diagnosis of individual needs; the identification of areas of strengths or weaknesses in the curriculum; and the redirection of content not mastered by students is possible with appropriate support. These improvements, however, can only be accomplished with efforts of all the stakeholders.

Promotion and graduation testing may also have unintended consequences for teachers. ... high-stakes testing is intended to raise teacher motivation and effectiveness, and there is evidence that with appropriate professional development, support, resources, and time teaching effectiveness can improve significantly (Elmore, 2000). There is already evidence, however that the negative publicity associated with poor test scores can lead experienced teachers to leave urban schools for the suburbs (See, *e.g.*, Lee, 1998). Plainly, efforts to improve low-performing urban schools - and to educate all children effectively - will be undermined if those schools lose strong teachers (Heubert, 2000, p.2).

Teachers, experts, and policymakers working together could make as their goal not the elimination of LEAP 21 testing but the identification of its strengths and weaknesses. After establishing its strengths and weaknesses then changes could be made to create a tool that is helpful to the teachers and beneficial to the students. This committee could also focus on states that are maintaining high achievement in the face of high stakes testing. Three of the highest rated states are Nebraska, Minnesota and Connecticut. “[These states] conduct performance-based assessments, instead of relying solely on high-stakes tests. In general, these assessments, scored by multiple teachers, look at samples of students' work, take into account behavioral observations and measure whether students have learned the material covered in course syllabi” (Smith, 2001, p. 4).

Another option that has led to more comprehensive evaluation and assessment without the high level of reported unintended consequences that high stakes test has engendered are the use of portfolios. The state of Kentucky has used portfolios successfully in response to “No

Child Left Behind” legislation (Borko & Elliott, 1999). Most states avoid the use of portfolios because the cost of assessing them is exorbitant. Additional research into the use of these methods of alternative assessments would be constructive in creating cheaper assessments that accurately reflect the students’ academic progress.

### Implications of the Study

An intention of this research is to generate suggestions that might guide school improvement and enhance student learning. My suggestions are meant to communicate how the findings of the study could inform various practical applications in education. These practical implications are discussed in the following section of this chapter.

My research presents policy makers with clear evidence that after eight years of LEAP 21 testing; teachers continue to communicate apprehension and frustrations with this assessment program. In this section, I will supply some proposals for altering the LEAP testing based on the teachers’ responses to the surveys and interviews conducted in this study.

One theme that was ascertained from the data is that the use and consequences of the SPS scores needs to be limited. Some teachers reported that the SPS scores could be used successfully to aid their teaching methods and progress student learning by providing information regarding student weaknesses. Nevertheless, the SPS scores were not perceived as being fair when employed to make comparisons between students, teachers, or schools for the various reasons stated in Chapter 5. Many of the teachers, regardless of past testing success of their classes, believed that their students were at a clear disadvantage. This disadvantage was the result of their students being held to the same standard as wealthier children who have more resources. Furthermore, some of the teachers were frustrated with how the SPS scores were calculated. The fact that the fourth grade was held to the same standard as the fourth grade the year before, while ignoring the differing capabilities of the two classes (cohorts) was noted. The majority of the



teachers stated that the current system was unfair. The teachers' statements suggest that policy makers should eliminate the current system or change the criteria to make it fairer. As I stated earlier, teachers are supported in their concerns that the current system is unfair. Comparing teachers and schools based on SPS scores is unjust because scoring success or failure could be attributed to other influences on students besides the teacher (Berk, 1988).

One possible solution to make the SPS scores more just would be to base the SPS scores on the socioeconomic status of the students "which is generally correlated with achievement" (Jones & Egley, 2004, p.25). Another possible solution would be to test students' knowledge at the beginning of the year and contrast these scores with their end-of-year scores. The result would be an assessment that more directly measures the effects of student learning and progress during that year.

Another main concern of the teachers was the use of a single test to accurately measure the students' academic progress. The LEAP test is actually offered twice: March (for all 4th graders) and in June (for 4th graders who failed in March). However the SPS score is only affected by the March test.

Another solution would be to offer the test more than once a year and average the scores. However, this solution would remove more time from instruction. A better solution may be to use an alternative method of assessment such as portfolios. "Portfolio assessments are a collection of student works and generally include a student's classroom work, revisions, assessments, and reflections on his or her learning. Some teachers have found that portfolios can positively impact their teaching methods and are essential to holding teachers accountable" (Jones & Egley, 2004, p.26). The state of Kentucky has used portfolios successfully in response to "No Child Left Behind" legislation (Borko & Elliott, 1999). Most states avoid the use of portfolios because the cost of assessing them is prohibitive. Additional research into the use of

these methods of alternative assessments would be constructive in creating cheaper assessments that accurately reflect the students' academic progress.

Of course another solution would be to eliminate high stakes testing altogether. As Chris Gallagher (2003) states:

Accountability systems must promote high-impact, not high-stakes, assessment. High-stakes accountability is really a form of accounting, in which the winners -- those with the most points (highest scores) -- are rewarded, and the losers -- those with the fewest points (lowest scores) -- are punished. Sadly, it is also based on a flawed learning theory. Any teacher who relies solely on extrinsic motivation -- the wielding of carrots and sticks -- is committing professional malpractice. But that's exactly what high-stakes accountability systems do (p.5).

Another concern that needs to be addressed is the disconnect between the amount of time to teach subjects and amount of time given to teach the material tested. The LEAP test forced teachers to compress into six months what they would usually teach over a nine-month period.

The GLEs need to be tailored to include fewer topics within each subject. This would allow time for greater depth of understanding. An easier solution may be to simply push the test back to later in the school year or only include subject matter that can be reasonably be taught before the test is administered. Making one of these changes would help teachers to administer their class time more successfully, resulting in increased student learning and less perceived pressure.

Another concern expressed by the teachers involved the incredible amount of pressure placed on fourth grade students and teachers due to the high stakes nature of the LEAP test. Policy makers would be well advised to provide a forum for teachers to voice their ideas about the effects of accountability in their classrooms. As stated before, recognizing teachers' perceptions are imperative, because they have the most direct effect on students' academic achievement. "What is needed is a policy shift that emphasizes inclusion of constituents" (DeBard & Kubow, 2002, p. 403).

The concerns stated by teachers regarding commitment would likely be diminished if a few of the recommendations provided in this section were realized. For example, employing the suggestions to decrease pressure on teachers and students would likely heighten teachers' commitment to the profession. Also, reducing teachers' perceived pressure may help reduce the possibility that they would engage in the practice of teaching to the test.

#### Recommendations for Future Research

Researchers that are interested in examining subjects analogous to the ones in this study can build upon the research by exploring several topics for deeper understanding. More studies are needed that utilize quantitative, qualitative, or mixed methods. These topics could include:

- organizing an inquiry similar to this one in Jefferson Parish public schools that include middle and high schools. This would determine if the patterns and themes discovered with teachers of elementary schools are unique or have a broader generalizability,
- replicating this research with a larger sample of elementary schools and an increased amount of teacher interviews would either confirm the patterns and themes ascertained in this study, or could demonstrate that within the larger population the results of this report are not representative of Jefferson Parish Public Schools. Also, the greater number of participants would allow a more reliable comparison of the demographics of the teachers and multi-level statistical analysis,
- studying the long term effects of the increasing pressure regarding teachers who are at schools whose SPS target scores keep increasing,
- replicating this research with a focus on the views of students regarding high stakes testing in addition to the teachers,

- continuing the study using a longitudinal approach, to see if there is a change in instruction, pressure, or commitment would help determine if the characteristics of high stakes testing is changing or constant,
- replicating this research with a focus on the effects of high stakes testing on instruction for both special and regular education.
- developing a framework for assessing the costs and benefits of LEAP 21 testing with a focus on alternative assessments.
- studying alternative assessments with a special emphasis on assisting low performing students

As this study demonstrated high stakes testing seems to have a negative impact on the perceptions and workloads of teachers, but there is little information about the effects on students themselves. “States still do not take into account the full costs of high stakes testing programs, and claims that testing alone can cause major educational improvements have not been proven” (Langenfeld, Thurlow, & Scott, 1997, p. 46).

Alternative assessments (such as portfolios) have potential, but it has not been established that these evaluations can be done effectively and reliably as a large-scale comparative assessment method. More studies are needed.

### Summary of Chapter 6

Five years ago, Amrein and Berliner's (2002) examination of data across 28 states, demonstrated that the focus on high-stakes testing had not increased student academic improvement on other standardized measures. This study supports many of Amrein and Berliner's (2002) findings that high stakes testing leads to outcomes such as a narrowing of the curriculum, teaching to the test, and teacher flight from public schools as well as from the profession. Capturing the perceptions of the participants of this study led to the realization that

an undue stress results from current LEAP 21 testing policies. This pressure affects both teachers and students. Studies have shown that pressure impacts not only test performance, but also motivation towards learning and teaching in general (Ryan & Grolnick, 1986).

In this chapter, a summary of the entire study was followed by a discussion concerning the results from the study. The practical policy implications that emanated from the discussion were then listed and future research topics were presented.

## REFERENCES

- Abington-Cooper, M.(2005). An evaluation of the LSU Agricultural Center's agriculture leadership development program, 1988-2004. Dissertation Abstracts International (UMI No. 155654).
- Abrams, L. M. & Madaus, O. F. (2004, November). The lessons of high-stakes testing. *Educational Leadership*, 61(3), 31-35.
- Abrams, L. M., Pedulla, J. J., & Madaus, G. E. (2003). Views from the classroom: Teachers' opinions of statewide testing programs. *Theory Into Practice*, 42(1). Retrieved March 3, 2007, from ERIC database.
- Adami-Bunyard, E., Gummow, M., & Milazzo-Licklider, N. (1998). Improving primary student motivation and achievement in mathematics. Retrieved March 3, 2007, from ERIC database.
- Afflerbach, P. (2005). National reading conference policy brief: High stakes testing and reading assessment. *Journal of Literacy Research*, 37(2), 151-62.
- Alamillo, L., & Viramontes, C. (2000). Reflections from the classroom: Teacher perspectives on the implementation of Proposition 227. *Bilingual Research Journal*, 24, 155-68.
- American Educational Research Association. (2000). AERA position statement concerning high stakes testing in Pre K-12 education. Retrieved November 22, 2005, from <http://www.aera.net/about/policy/stakes.htm>.
- American Federation of Teachers (AFT), AFL-CIO. (2001). Making standards matter. Retrieved November 21, 2005 from <http://www.aft.org/>.
- American Psychological Association. (2001). Appropriate use of high-stakes testing in our nation's schools. Retrieved April 8, 2008, from the World Wide Web: <http://www.apa.org/pubinfo/testing.html>.
- Amrein, A. L. & Berliner, D. C. (2003). The testing divide: New research on the intended and unintended impact of high-stakes testing. *Peer Review*, 31-32.
- Amrein, A. L. & Berliner, D. C. (2002). High-stakes testing, uncertainty, and student learning. *Education Policy Analysis Archives*, 10(18). Retrieved November 25, 2005 from <http://epaa.asu.edu/epaa/v10n18/>
- Appropriate use of high-stakes testing in our nation's schools. Retrieved on November 22, 2005 from <http://www.apa.org/pubinfo/testing.html>.
- Bandura, A. (1986). *Social Foundations of Thought and Action: a Social Cognitive Theory*. Englewood Cliffs, NJ: Prentiss-Hall.

- 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*, 5, 384-97.
- Berk, R. A. (1988). Fifty reasons why student achievement gain does not mean teacher effectiveness. *Journal of Personnel Evaluation in Education*, 1(4), 345-64.
- Bishop, J. H., & Mane, F. (2001). The impacts of minimum competency exams graduation requirements on high school graduation, college attendance, and early labor market success. *Labor Economics*, 8(2), 203-22.
- Bobbie, E. (1990). *Survey research methods* (2<sup>nd</sup> ed.). Belmont, CA: Wadsworth.
- Bogdan, R. C. & Biklen, S. K. (1998). *Qualitative research for education: An introduction to theories and methods* (4<sup>th</sup> ed.). Boston: Allyn and Bacon.
- Bond, L. A., & King, D. (1995). *State high school graduation testing: Status and recommendations*. Oakbrook, IL: North Central Regional Educational Laboratory.
- Borko, H., & Elliot, R. (1999). Hands-on pedagogy versus hands off accountability: Tensions between competing commitments for exemplary math teachers in Kentucky. *Phi Delta Kappan*, 80, 394-400.
- Buck, J., Evans, D., & Nguyen, A. (2006). Accountability and high stakes testing: Texas examinations of education standards-principal. *The Lamar University Electronic Journal of Student Research*, 3, 1-4.
- Bunting, C. (1999). School reform: Does it really matter? *The Clearing House*, 72(4), 45-55.
- Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56, 81-105.
- Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research*. Chicago: Rand McNally.
- Cankoy, O., & Tut, M. A. (2005). High- stakes testing and mathematics performance of fourth graders in North Cyprus. *Journal of Educational Research*, 98(4), 234-43.
- Carpenter, S. (2001). The high stakes of educational testing. *Monitor on Psychology*, 32(5), 1-2.
- Caruso, D. B. (2005, July 5). High stakes tests spur pupils to pass. *Seattle Post-Intelligencer*, p. A1.
- Cimbricz, S. (2002). State-mandated testing and teachers' belief and practice. *Education Policy Analysis Archives*, 10(2). Retrieved December 24, 2006, from <http://epaa.asu.edu/epaa/v10n2.html>.
- Clark, N. B. (2005). Toward a construct validation of the Louisiana school analysis model instructional staff questionnaire. *Dissertation Abstracts International* (UMI No. 3166909).

- Coleman, J. S., Campbell, E., Hobson, C., McPartland, J., Mood, A., Weinfeld, F., & York, R. (1966). *Equality of educational opportunity*. Washington, DC: U. S. Government Printing Office.
- Corbett, H. D., & Wilson, B. L. (1991). *Testing, reform, and rebellion*. Norwood, NJ: Ablex.
- Cortiella, C. (2004). Implications of high-stakes testing for students with learning disabilities. Retrieved March 6, 2007, from SchwabLearning.org: <http://www.schwablearning.org/articles.aspx?r=846>.
- Creswell, J. W. (1994). *Research design: Qualitative and quantitative approaches*. Thousand Oaks, CA: Sage.
- Creswell, J. W. (2002). *Educational research: Planning, conducting, and evaluating quantitative and qualitative approaches*. Upper Saddle River, NJ: Merrill.
- Creswell, J. W., Clark, V. L. P., Gutmann, M. L. & Hanson, W. E. (2003). Advanced mixed methods research designs. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (pp. 209 – 40). Thousand Oaks, CA: Sage.
- Creswell, J. W. (2003). *Research design: Qualitative, quantitative, and mixed methods approaches* (2<sup>nd</sup> ed.). Thousand Oaks, CA: Sage.
- Creswell, J. W., & Clark, V. L. P. (2007). *Designing and conducting mixed methods research*. Thousand Oaks, CA: Sage.
- Crone, L. J., Franklin, b. J., Caldas, S. J., Ducote, J. M., & Killebrew, C. (1992). The use of norm-referenced and/or criterion-referenced tests as indicators of school effectiveness. *Louisiana Educational Research Journal*, 56 – 77.
- Darling-Hammond, L., & Falk, B. (1997). Policy for authentic assessment. In *Assessment for Equity and Inclusion: Embracing All Our Children*. New York: Routledge.
- Davis, A. (2006). Theories used in IS research: Social cognitive theory. Boone, NC: Appalachian State University, Theories Used in IS Research. Retrieved June 9, 2007, from <http://www.istheory.yorku.ca/socialcognitivetheory.htm>.
- DeBard, R., & Kubow, P. K. (July, 2002). From compliance to commitment: The need for constituent discourse in implementing testing policy *Educational Policy* 16 (3), 387-405
- Debra P. v. Turlington, 474 F. Supp. 244 (M. D. Fla. 1979); aff'd in part, rev'd in part, 644 F. 2d 397 (5th Cir 1981).
- DeCesare, D. (2002). How high are the stakes in high-stakes testing? *Principal: The standardized curriculum*, 81, 3, 10-12.



- Dede, C. (2004). Enabling distributed learning communities via emerging technologies. *The Journal*, 32(2), 12-22.
- Denzin, N. K., & Lincoln, Y. S. (1994). Introduction: Entering the field of qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 1-18). Thousand Oaks, CA: Sage.
- Dörnyei, Z. (2001). *Teaching and researching motivation*. Essex, England: Pearson Education.
- Dorn, R. I., Douglass, J., Ekiss, G. O., Trapido-Lurie, B., Comeaux, M., Mings, R., Eden, R., Davis, C., Hinde, E., & Ramakrishna, B. (2005) Learning geography promotes learning math: Results and implications of Arizona's GeoMath Grade K-8 Program. *Journal of Geography*, 104(4), 151-9.
- Ed.gov. (2003). Fact sheet on the major provisions of the conference report to H.R. 1, the No Child Left Behind Act. Retrieved November 30, 2005, from <http://www.ed.gov.org>.
- Eisenhart, M. A., & Howe, K. R. (1990, April). Validity in qualitative research. Paper presented at the annual meeting of the American Educational Research Association, Boston, MA.
- Elmore, R. (2000). *Building a new structure for school leadership*. Washington, DC: The Albert Shanker Institute.
- Erickson, F. (1986). Qualitative methods in research on teaching. In M. Wittrock (Ed.) *Handbook of Research on Teaching*, 3, 119-161. New York: MacMillan.
- Evans, L. (1998). *Teacher morale, job satisfaction and motivation*. London: Paul Chapman.
- FairTest (n.d.). Initial FairTest analysis of ESEA as passed by congress, Dec. 2001. Retrieved December 15, 2006, from <http://www.fairtest.org/natest/ESEA.html>.
- Finely, S.J. (1999). The progress of education in Louisiana, Southwest Educational Development Laboratory. Retrieved November 23, 2005, from <http://www.sedl.org/pitl/pic/states/la/welcome.html>.
- Firestone, W. A., Mayrowetz, D., & Fairman, J. (1998). Performance-based assessment and instructional change: The effects of testing in Maine and Maryland. *Educational Evaluation and Policy Analysis*, 20(2), 95-113.
- Firestone, W. A., Monfils, L., Camilli, G., Schorr, R., Hicks, J. & Mayrowetz, D. (2001). The ambiguity of test preparation: A multi-method analysis in one state. Paper presented at the meeting of the American Educational Research Association, Seattle, Washington, April.
- Fink, A. (1995). *How to ask survey questions*. Thousand Oaks, CA: SAGE Publications.
- Franklin, B.J., Pernici, S.C., & Yuan, X.S. (2005). Grade level retention rates in Louisiana public schools 1997-1998 to 2000-2005. ERIC document No. Ed 455301.

- Freeman, J. A. (1997). A methodological examination of naturally occurring school improvement in Louisiana schools. Unpublished doctoral dissertation, Louisiana State University, Baton Rouge, LA. (UMI No. 9736015).
- Gallagher, C. (2003). Turning the accountability tables: Ten progressive lessons from one 'backward' state. *Phi Delta Kappan*, 85, 5, 352-60.
- Getzels, J.W. & Guba, E.G. (1957). Social behavior and the administrative process. *The School Review*, 65, 423-41.
- Glanz, K., Rimer, B.K. & Lewis, F.M. (2002). *Health behavior and health education. Theory, Research and Practice*. San Francisco: Wiley & Sons.
- Glaser, B. G. (1978). *Theoretical sensitivity*. Mill Valley, CA: Sociology Press.
- Goertz, M., & Duffy, M. (2003). Mapping the landscape of high-stakes testing and accountability programs. *Theory Into Practice*, 42, 4-11.
- Goertz, M.E., Floden, R.E., & O'Day, J. (1995) *Studies of Education Reform: Systemic Reform, Volume 1: Findings and Conclusions*. United States Department of Education, Washington, D.C., 1-164.
- Gould, S. J. (1996). *The mismeasure of man*. New York: W. W. Norton Company, Inc.
- Goldberg, M. F. (2004). The test mess. *Education Digest: Essential Readings Condensed for Quick Review*, 69, 8, 8-15.
- Goldhaber, D. (2002). What might go wrong with the accountability measures of the No Child Left Behind Act? Retrieved December 11, 2006, from The Urban Institute Web Site: <http://www.edexcellence.net/NCLBconference/Goldhaber.pdf>.
- Grant, M. M. & Hill, J. R. (2006). Implementing student-centered pedagogy within high stakes testing. In R. Lambert & C. McCarthy's (Eds.), *Understanding Teacher Stress in an Age of Accountability* (pp. 19-42), Greenwich, CT: Information Age Publishing.
- Greene, J. C., Caracelli, V. J., & Graham, W. F. (1989). Toward a conceptual framework for mixed-method evaluation designs. *Educational Evaluation and Policy Analysis*, 11, 255-74.
- Guglielmi, R. S., & Tatrow, K. (1998). Occupational stress, burnout, and health in teachers: A methodological and theoretical analysis. *Review of Educational Research*, 68(1), 61-100.
- Hair, D., Kraft, B., & Allen, A. (2001). National Staff Development Council Project ADVANCE mini-grant: Louisiana Staff Development Council's end of grant report. Baton Rouge, LA: Louisiana Staff Development Council.
- Hair, J., Tatham, R., Anderson, R., & Black, W. (1998). *Multivariate data analysis*. (5th ed.). New York: Prentice Hall

- Halsey, A. H. (1995). *Decline of Donnish dominion*. Oxford, England: Clarendon Press.
- Hamilton, L. S., & Koretz, D. M. (2002). Tests and their use in test-based accountability systems. In L. S. Hamilton, B. M. Stecher, & S. P. Klein, (Eds.) *Making sense of test-based accountability in education* (pp. 13-49). Santa Monica, CA: RAND.
- Hamilton, L., & Stecher, B., (2004). Responding effectively to test-based accountability. *Phi Delta Kappan*, 85, 578-83.
- Haney, W. (2000). The myth of the Texas miracle in education. *Education Policy Analysis Archives*, 8(41). Retrieved December 24, 2006, from <http://epaa.asu.edu/epaa/v8n41/>.
- Hatcher, L. (1994). *A step by step approach to using the SAS(R) system for factor analysis and structural equation modeling*. Cary, NC: SAS Institute, Inc.
- Hauser, R. M., Pager, D. I. & Simmons, S. J. (2000, August). Race-ethnicity, social background, and grade retention. Presented at the annual meeting of the American Sociological Association, Washington, D.C.
- Hendrie, C. (1996, October 9). 109 Chicago schools put on academic probation. *Education Week*, 16(6), 1-2.
- Herman, J., & Golan, S. (1991). *Effects of standardized testing on teachers and learning: Another look* (CSE Tech. Rep. No. 334). Los Angeles: university of California, Center for the Study of Evaluation.
- Heubert, J. P., & Hauser, R. M. (Eds.). (1999). *High stakes: Testing for tracking, promotion, and graduation*. National Research Council, Washington, DC: National Academy Press
- Heubert, J. (2000). Potential benefits and risks for minority students, English-language learners, and students with disabilities. *Poverty and Race* 9 (5): 1-2, 5-7. Washington, DC: Poverty and Race Research Action Council.
- High stakes testing. (2003, January 7). *The St. Petersburg Times*, p. A1.
- Hochschild, H., & Scott, B. (1998). Trends: Governance and reform of public education in the United States. *Public Opinion Quarterly*, 62(1), 79-120.
- Hoepfl, M. C. (1997). Choosing qualitative research: A primer for technological education researchers. *Journal of Technology Education*, 9(1), 47-63.
- Hoffman, J. V., Assaf, L. C., & Paris, S. G. (2001). High-stakes testing in reading: Today in Texas, tomorrow? *The Reading Teacher*, 54, 482-92.
- Hombo, C. (2003). NAEP and no child left behind: Technical challenges and practical solutions. *Theory Into Practice*, 42, 59-65.

- Hoy, W. K. & Miskel, C. G. (1996). *Educational Administration: Theory, Research and Practice* (5th ed.). New York: McGraw-Hill, Inc.
- Hursh, D. (2005). The growth of high stakes testing in the USA: Accountability, markets, and the decline in educational equality. *British Educational Research Journal*, 31(5), 605-22.
- Ingersoll, R. (2001). Teacher turnover and teacher shortages: An organizational analysis. *American Educational Research Journal*, 38(3), 499-534.
- Jacob, B. A. (2005). Accountability, incentives, and behavior: The impact of high-stakes testing in the Chicago Public Schools. *Journal of Public Economics*, 89(5), 761-96.
- Jencks, C. S., Smith, M., Acland, H., Bane, M. J., Cohen, D., Ginter, H., Heyns, B., & Michelson, S. (1972). *Inequality: A reassessment of the effect of family and schooling in America*. New York: Basic Books.
- Johnson, B., & Turner, L.A. (2003). Data collection strategies in mixed methods research. In A. Tashakkori & C. Teddlie's *Handbook of mixed methods in social & behavioral research*, Thousand Oaks, CA: Sage.
- Johnson, D., & Johnson, B. (2006). *High stakes: Poverty, testing, and failure in American schools*. New York: Rowman & Littlefield.
- Johnson, R. A. & Wichern, D. W. (1992). *Applied multivariate statistical analysis*. (3rd ed.). Englewood Cliffs: Prentice-Hall, Inc.
- Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher*, 33(7), 14-26.
- Jones, B. D. & Egley, R. J. (2004, August 9). Voices from the frontlines: Teachers' perceptions of high-stakes testing, *Education Policy Analysis Archives*, 12(39). Retrieved January 11, 2008 from <http://epaa.asu.edu/epaa/v12n39/>.
- Jones, B. D., & Johnston, A. F. (2002, April). The effects of high stakes testing on instructional practices. Paper presented at the meeting of the American Educational Research Association, New Orleans, LA.
- Jones, M. G., Jones, B. D., & Hargrove, T. Y. (2003). *The unintended consequences of high stakes testing*. Lanham, MD: Rowman & Littlefield.
- Jones, M. G., Jones, B. D., Hardin, B. H., Chapman, L., Yarbrough, T., & Davis, M. (1999). The impact of high stakes testing on teachers and students in North Carolina, *Phi Delta Kappan*, 81(3), 193-203.
- Kay, A. C., Jimenez, M. C. & Jost, J. T. (2002). Sour grapes, sweet lemons, and the anticipatory rationalization of the status quo. *Personality and Social Psychology Bulletin*, 28(9), pp. 1300-12.

- Kemper, E., Stringfield, S., Teddlie, C. (2003). Mixed methods sampling strategies in social science research. In A. Tashakkori & C. Teddlie's Handbook of mixed methods in social & behavioral research, Thousand Oaks, CA: Sage.
- Kellaghan, T., Madaus, G., & Airasian, P. (1980). The effects of standardized testing. Educational Research Centre: St. Patrick's College, Dublin, Ireland, Boston College: Chestnut Hill, MA.
- Kiley, K., & Henry, T. (2001, December 17). Will No Child Be Left Behind? USA Today, pp. D4.
- Kiziltepe, Z. (2006). Sources of teacher demotivation. In R. Lambert & C. McCarthy's (Eds.), Understanding Teacher Stress in an Age of Accountability (pp. 145-62), Greenwich, CT: Information Age Publishing.
- Klein, S., Hamilton, L., McCaffrey, D., & Stecher, B. (2000). What do test scores in Texas tell us? (IP-202) Santa Monica, CA: RAND.
- Kohn, A. (2000, September 27). Standardized testing and its victims. Education Week, 20(4), 46-47, 60.
- Koretz, D. (1996). The assessment of students with disabilities in Kentucky (CSE Technical Report No. 431), Los Angeles: CRESST/Rand Institute on Educational Training.
- Koretz, D. M., Barron, S., Mitchell, K. J., & Stecher, B. M. (1996). Perceived effects of the Kentucky Instructional Results Information System (KIRIS). Santa Monica, CA: RAND.
- Krueger, R. A., & Casey, M. A. (2000). Focus group interviews: A practical guide for applied research. (3rd ed.). Thousand Oaks, CA: Sage.
- Kymes, N. (2004). The No Child Left Behind Act: A look at provisions, philosophies, and compromises. Journal of Industrial Teacher Education. 41(2), p. 1-8.
- Lange, K. (2002). Mathematical and statistical methods for genetic analysis (2<sup>nd</sup> ed.). New York: Springer-Verlag.
- Langenfeld, K. L., Thurlow, M. L., & Scott, D. L. (1997). High-stakes testing for students: Unanswered questions and implications for students with disabilities (Synthesis Report 26). Minneapolis: University of Minnesota, National Center on Educational Outcomes.
- LaSIP Overview. Retrieved, November 22, 2005 from <http://www.lasip.state.la.us> currently unavailable.
- Lazear, E. P. (2006). Speeding, terrorism, and teaching to the test. Quarterly Journal of Economics, 121(3), 1029-61.
- Lincoln, Y. S., & Guba, E. G. (1985). Naturalistic inquiry. Beverly Hills, CA: Sage.

- Lofland, J., & Lofland, L. H. (1984). *Analyzing social settings*. Belmont, CA: Wadsworth Publishing Company, Inc.
- Louisiana Department of Education. (2000). *Louisiana state education progress report 1999-2000*. Baton Rouge, LA: Author.
- Louisiana Department of Education (2001). *LEAP21*. Baton Rouge, LA: Author.
- Louisiana Department of Education. (2003). *LEAP for the 21st century high stakes testing policy*. Retrieved November 22, 2005 from <http://www.doe.state.la.us/lde/uploads/1703.pdf>.
- Louisiana Department of Education. (2004). *Section 1: The Louisiana educational assessment program LEAP21/GEE 2002-2003 annual report*. Baton Rouge, LA: Author.
- Louisiana Department of Education. (2005). *Louisiana state education progress report 2004-2005*. Baton Rouge, LA: Author.
- Louisiana Department of Education. (2006). *Principal's Report Card*. Baton Rouge, LA: Author.
- Louisiana Department of Education. (2007). *Bulletin 111: Louisiana school, district, and state accountability policy*. Retrieved March 23, 2007, from <http://www.louisianaschools.net/lde/bese/home.html>.
- Louisiana District and School Accountability Advisory Commission. (1998, March 18). *Recommendations for Louisiana's public education accountability system*. Baton Rouge: Author.
- Madaus, G. (1998). *The distortion of teaching and testing: High-stakes testing and instruction*, *Peabody Journal of Education*, 65, 29-46.
- Madaus, G., West, M., Harmon, M., Lomax, R., & Viator, K. (1992). *The influence of testing on teaching math and science in grades 4-12*. Boston: Center for the Study of Testing, Evaluation, and Educational Policy, Boston College.
- Make-or-break exams grow, but big study doubts value. (2002, December 28). *The New York Times*, p. A1.
- Mancuso, M. L. (2004) *High stakes testing in Louisiana: An analysis of the disparate impact on black and white eighth-grade students and the perspectives of parents*. Dissertation Abstracts International (UMI No. 3151836).
- Manzo, K. K. (1996, October 30). *Philadelphia plan links student achievement, teacher pay*, *Education Week*, 16(9), 2-5.
- Maslach, D., Schaufeli, W. B., & Leiter, M. P. (2001). *Job burnout*. *Annual Review of Psychology*, 52, 397-422.

- Mathison, S., & Freeman, M. (2006). Teacher stress and high stakes testing. How using one measure of academic success leads to multiple teacher stressors in R. G. Lambert & C. J. McCarthy (Eds.) *Understanding Teacher Stress in an Age of Accountability*. Greenwich, CT: Information Age Publishing.
- Maxcy, S. J. (2003). Pragmatic threads in mixed methods research in the social sciences: The search for multiple modes of inquiry and the end of the philosophy of formalism. In A. Tashakkori & C. Teddlie's (Eds.), *Handbook of Mixed Methods in Social & Behavioral Research*, Thousand Oaks, CA: Sage.
- Maxwell, J. A. (1996). *Qualitative research design: An interactive approach*. Thousand Oaks, CA: Sage.
- McDonnell, L. M., McLaughlin, M. J., & Morrison, P. (1997). *Educating one and all: Students with disabilities and standards-based reform*. Washington, D.C: National Academy Press.
- McGuire, W. J., & McGuire, C. V. (1991). The content, structure, and operation of thought systems. In R. S. Wyer, Jr. & T. K. Srull (Eds.), *Advances in social cognition*. Hillsdale, NJ: Lawrence Erlbaum.
- McMillan, J., Myran, S., & Workman, D. (1999, April). The impact of mandated statewide testing on teachers' classroom assessment and instructional practices. Paper presented at the annual meeting of the American Educational Research Association, Montreal, Quebec, Canada.
- McNeil, L. M. (2000). Creating new inequalities: Contradictions of reform. *Phi Delta Kappan*, 81, 729-34.
- McNeil, L., & Valenzuela, A. (2000). The harmful impact of the TASS system of testing in Texas: Beneath the accountability rhetoric. In G. Orfield & M. Kornhaber (Eds.), *Raising Standards or Raising Barriers: Inequality and High Stakes Testing in Public Education* (pp. 127-150). Cambridge, MA: Harvard Civil Rights Project.
- Mehrens, W. A. (1998). Consequences of assessment: What is the evidence? Vice presidential address for division D, annual meeting of the American Educational Research Association, San Diego, CA.
- Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook* (2<sup>nd</sup> ed.). Thousand Oaks, CA: Sage.
- Miller, S. (2003). Impact of mixed methods and design on inference quality. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (pp. 423 – 56). Thousand Oaks, CA: Sage.
- Morgan, D. (1998). Practical strategies for combining qualitative and quantitative methods: Applications to health research. *Qualitative Health Research*, 8, 362-76.

- Morse, J. M. (1991). Approaches to qualitative-quantitative methodological triangulation. *Nursing Research*, 40(2), 120-23.
- National Testing. (No Child Left Behind). Retrieved November 22, 2005, from <http://www.fairtest.org>.
- Natriello, G., & McDill, E. L. (1999). Title I: From funding mechanism to educational program. The Harvard Civil Rights Project, *Hard work for good schools: facts not fads in Title I reform*. Cambridge, MA: Harvard University.
- Nieto, S. (2003). Challenging current notions of “highly qualified teachers” through work in a teachers’ inquiry group. *Journal of Teacher Education*, 54(5), 386-98.
- No Child Left Behind Act of 2001, Pub. L. No. 107-220, 115 Stat. 1425 (2002).
- O'Neill, P. T. (2003). High stakes testing law and litigation. *Brigham Young University Education and Law Journal*, 2003, pp. 623-62.
- Onwuegbuzie, A. J. (2000, November). Validity and qualitative research: An oxymoron? Paper presented at the annual meeting of the Association for the Advancement of Educational Research, Ponte Verda, FL.
- Onwuegbuzie, A. J., & Teddlie, C. (2003). A framework for analyzing data in mixed methods research. In A. Tashakkori & C. Teddlie's *Handbook of mixed methods in social & behavioral research*, Thousand Oaks, CA: Sage.
- Organization for Economic Cooperation and Development. (1989). *The condition of teaching: General report*. In M. Fullan & A. Horgreaves (Eds.), *Teacher development and educational change*. London: The Falmer Press.
- Pane, J. F., McCaffrey, Tharp-Taylor, S. Asmus, G. J., & Stokes, B. R. (2006). Student displacement in Louisiana after the hurricanes of 2005: Experiences of public schools and their students. Santa Monica, CA: RAND Gulf States Policy Institute, Retrieved March 28, 2007 from [http://www.rand.org/pubs/technical\\_reports/2006/RAND\\_TR430.pdf](http://www.rand.org/pubs/technical_reports/2006/RAND_TR430.pdf).
- Paterno, J. (2004). Important definitions. *Building Cathedrals*, Retrieved November 22, 2005 from <http://www.angelfire.com/wa2/buildingcathedrals/measuringsuccess.html>.
- Patton, M. Q. (1990). *Qualitative evaluation and research methods* (2<sup>nd</sup> ed.). Newbury Park, CA: Sage.
- Patton, M. Q. (1998). Discovering process use. *Evaluation*, 4, 225-33.
- Patton, M. Q. (2002). *Qualitative evaluation and research methods* (3rd ed.). Newbury Park, CA: Sage.



- Pedulla, J. J., Abrams, L. M., Madaus, G. F., Russell, M. K., Ramos, M. A., & Miao, J. (2003). Perceived effects of state-mandated testing programs on teaching and learning: Findings from a national survey of teachers. Boston, MA: Boston College, National Board of Educational Testing and Public Policy. Retrieved December 20, 2006, from <http://www.bc.edu/research/nbetpp/statements/nbr2.pdf>.
- Phelps, R. P. (2003). Kill the messenger: The war on standardized testing. Somerset, N. J.: Transaction Publishers.
- Pringle, R. M., & Martin, S. C. (2005). The potential impacts of upcoming high-stakes testing on the teaching of science in elementary classrooms. *Research in Science Education*, 35(2), 347-61.
- Public Affairs Research Council. (2002, February). Close-Up on education accountability. Retrieved February 11, 2007, from <http://www.la-par.org/Publications/PDF/Closeup2.pdf>.
- Quick, J. C., Quick, J. D., Nelson, D. L., & Hurrell, J. J., Jr. (1997). Preventive stress management in organizations. Washington, DC: American Psychological Association.
- Randall, K. (2001). Growing opposition to "high-stakes" testing in US schools. Retrieved January 11, 2008, from <http://www.wsws.org/articles/2001/jul2001/test-j25.shtml>.
- Ravitch, D. (2000). Left back. A century of failed school reform. New York: Simon and Schuster.
- Reardon, S. (1996, April). Eighth grade minimum competency testing and early high school dropout patterns. Paper presented at the annual meeting of the American Educational Research Association, New York.
- Robelen, E. W. (2000, May 24). La. set to retain 4th, 8th graders based on state exams. *Education Week*, 19 (37) 21-24.
- Roderick, M., Jacob, B. A., & Bryk, A. S. (2002). The impact of high stakes testing in Chicago on student achievement in promotional gate grades. *Educational Evaluation and Policy Analysis*, 24(4), 333-58.
- Rose, E. (2002). A special issue on school-university partnerships in special education. *Remedial and Special Education*, 23, 322.
- Rossmann, G. B., & Wilson, B. L. (1985). Numbers and words: Combining quantitative and qualitative methods in a single large-scale evaluation study. *Evaluation Review*, 9(5), 627-43.
- Russek, B. E., & Weinberg, S. L. (1993). Mixed methods of a study of implementation of technology-based materials in elementary classroom. *Evaluation and Program Planning*, 16, 131-42.

- Ryan, R. M., & Grolnick, W. S. (1986). Origin and pawns in the classroom: Self-report and projective assessments of individual differences in children's perceptions. *Journal of Personality and Social Psychology*, 50, 550-558.
- Salkind, N.J. (2004). *Statistics for people that hate statistics*. Thousand Oaks, CA: Sage.
- Salvaterra, M., & Adams, D. (1995). Departing from tradition: Two schools' stories. *Educational Leadership*, 53(3), 32-35.
- Sanders, W., & Horn, S. (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, 247.
- Schultz, B. D., & Oyler, C. (2006). We make this road as we walk together: Sharing teacher authority in a social action curriculum project. *Curriculum Inquiry*, 36(4), 423-51.
- Sclafani, S. (2002). No child left behind. *Issues in Science & Technology*, 19, 43-47.
- Seidman, I. E. (1991). *Interviewing as qualitative research*. New York: Teachers College Press.
- Seuss, Dr., Prelutsky, J., & Smith, L. (1998). *Hooray for diffendoofer day!* New York: Alfred A. Knopf.
- Shepard, L. A., & Dougherty, K. (1991, April). Effects of high stakes testing on instruction. Paper presented at the annual meeting of the American Educational Research Association, Chicago.
- Simon, D. B. (2008, January 17). School bonus plan gets ok. *Times-Picayune*, p. B1.
- Sirotnik, K. (2002). Promoting responsible accountability in schools and education. *Phi Delta Kappan*, 83(9), 662-73.
- Smith, D. (2001). Is too much riding on high-stakes tests? *Monitor On Psychology*, 32(11), 1-5.
- Smith, E. (2005). Raising standards in American schools: The case of No Child Left Behind. *Journal of Education Policy*, 20(4), 507-24.
- Smith, M. (1991). Put to the test: The effects of external testing on teachers. *Educational Researcher*, 20, 8-11.
- Smith, M. L., & Rottenberg, C. (1991). Unintended consequences of external testing in elementary schools. *Educational Measurement: Issues and Practice*, 10, 7-11.
- Smith, M. L., Edelsky, C., Draper, K., Rottenberg, C., & Cheuland, M. (1998). *The role of testing in elementary schools*. Los Angeles, CA: Center for Research on Educational Standards and Student Tests, Graduate School of Education, UCLA.

- Smith, M. L., Edelsky, C., Draper, K., Rottenberg, C., & Cheuland, M. (1989). The role of testing in elementary schools. Tempe, AZ: Arizona State University, Center for Research on Evaluation, Standards, and Student Testing (Grant No. OERI-G-86-0003).
- Smith, M., Nobel, A., Heinecke, W., Seck, M., Parish, C., Cabay, M. et al. (1997). Reforming schools by reforming assessment: Consequences of the Arizona student assessment program (ASAP): Equity and teacher capacity building. (CSE Technical Report 425). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing.
- Smith, M. S., & O'Day, J. A. (1993). Systemic school reform. In *Politics of Education Association Yearbook* (pp. 233-67). Taylor & Francis, Ltd.
- Spickerman, W. R. (1970). A study of the relationships between attitudes toward mathematics and some selected pupil characteristics in a Kentucky high school. (Doctoral dissertation, University of Kentucky, 1965). *Dissertation Abstracts International*, 30, 2733A.
- Stake, R. E. (1978, February). The case study method in social inquiry. *Educational Researcher*, 7(2), 5-8.
- Stecher, B., Barron, S., Chun, T., & Ross, K. (2000). The effect of Washington state education reform on schools and classrooms (CSE Technical Report 525). Los Angeles: National Center for Research on Evaluation, Standards, and Student Testing (CRESST).
- Strahan, D. (2003). General patterns and particular pictures: Lessons learned from reports from "beating the odds" schools. *Journal of Curriculum and Supervision*, 18, 296-305.
- Strike, K. A., Haller, E. J., & Soltis, J. F. (1998). *The ethics of school administration*. New York: Teacher College Press.
- Teddlie, C., Tashakkori, A., & Johnson, B. (2008, in press). Emergent techniques in the gathering and analysis of mixed methods data. In Leavy & Hesse-Biber (Eds), *Handbook of Emergent Methods*. New York: Guilford Press.
- Tashakkori, A. & Teddlie, C. (1998). *Mixed methodology: Combining the qualitative and quantitative approaches* (Applied Social Research Methods, No. 46). Thousand Oaks, CA: Sage.
- Tashakkori, A., & Teddlie, C. (2003). *Handbook of mixed methods in social & behavioral research*, Thousand Oaks, CA: Sage.
- Taylor, G., Shepard, L., Kinner, F., & Rosenthal, J. (2003). A survey of teachers' perspectives on high-stakes testing in Colorado: What gets taught, what gets lost (CSE Technical Report 588). Los Angeles: CRESST.
- Teddlie, C. (1994). Using context variables in school effectiveness research. In Reynolds, D. et al., *Advances in School Effectiveness Research and Practices*, Oxford: Pergamon, pp. 85-110.

- Teddlie, C. & Reynolds, D. (2000). *The international handbook of school effectiveness research*, New York: Falmer Press.
- The College Board is located at <http://www.collegeboard.com/splash>.
- Thomas, R. M. (2005). *High-stakes testing: Coping with collateral damage*. Mahwah, NH: Lawrence Erlbaum.
- Thompson, B. (1998). Statistical significance and effect size reporting: Portrait of a possible future. *Research in the Schools*, 5(2), 33-8.
- Tolbert, L. (2003). *The impact of Louisiana's school and district accountability system on students' performance on the state mandated criterion referenced test*. Dissertation Abstracts International (URN No. 0123103-194128).
- Tuerk, P. W. (2005). Research in the high-stakes era. *Psychological Science*, 16(6), 419-25.
- United States Department of Education, Office of Civil Rights, *supra* note 6, Introductory Letter, at 2. High-stakes tests and students with disabilities. Retrieved November 21, 2005 from [http://www.bc.edu/bc\\_org/avp/law/lwsch/journals/bclaw/41\\_5/03.txt.htm](http://www.bc.edu/bc_org/avp/law/lwsch/journals/bclaw/41_5/03.txt.htm).
- Urduan, T., & Paris, S. (1994). Teachers' perceptions of standardized achievement tests. *Educational Policy*, 8(2), 137-56.
- Waber, D. P., Gerber, E. B., Turcios, V. Y., Wagner, E. R., & Forbes, P. W. (2006). Executive functions and performance on high-stakes testing in children from urban schools. *Developmental Neuropsychology*, 29(3), 459-77.
- Wallace, B. & Graves, W. (1995). *Poisoned Apple*. New York: St. Martin's Press.
- Webb, P. T. (2006). The stress of accountability: Teachers as policy brokers in a poverty school. In R. Lambert & C. McCarthy's (Eds.), *Understanding Teacher Stress in an Age of Accountability* (pp. 1-18), Greenwich, CT: Information Age Publishing.
- Willms, J. D. (1986). Social class segregation and its relationship to pupil's examination results in Scotland. *American Sociological Review*, 51(2), 224-41.
- Wright, W. E. (2002, June 5). The effects of high stakes testing in an inner-city elementary school: The curriculum, the teachers, and the English language learners. *Current Issues in Education* [on-line], 5 (5). Available: <http://cie.ed.asu.edu/volume5/number5/>.
- Yarbrough, T. L. (1999). *Teacher perceptions of the North Carolina ABC program and the relationship to classroom practice*. Ph.D. Dissertation, University of North Carolina at Chapel Hill.
- Yeh, S. S. (2005). Limiting the unintended consequences of high stakes testing. *Education Policy Analysis Archives*, 13, 1-23.

Young, D. J. (2000). Teacher morale in Western Australia: A multilevel model. *Learning Environments Research*, 3, 159-77.

**APPENDIX A: STATE BY STATE DATA CONCERNING HIGH STAKES TESTING**

Table A.1:

Exit Examinations

	State exit or end-of-course exams are based on 10th grade standards or higher	State requires remediation for students failing exit, end-of-course, or promotion exams	State has a non-standard diplomas or tiered diploma system
Alabama	Yes	Yes	No
Alaska	Yes	No	Yes
Arizona	Yes	No	No
Arkansas	No	No	No
California	No	Yes	No
Colorado	No	No	No
Connecticut	No	No	No
Delaware	No	Yes	No
District of Columbia	No	No	No
Florida	Yes	Yes	Yes
Georgia	Yes	Yes	Yes
Hawaii	No	No	No
Idaho	Yes	No	No
Illinois	No	No	No
Indiana	No	Yes	No
Iowa	No	No	No
Kansas	No	No	No
Kentucky	No	No	No
Louisiana	Yes	Yes	Yes
Maine	No	No	No
Maryland	Yes	Yes	No
Massachusetts	Yes	No	No
Michigan	No	Yes	No
Minnesota	No	Yes	No
Mississippi	Yes	No	No
Missouri	No	Yes	No
Montana	No	No	No
Nebraska	No	No	No
Nevada	Yes	Yes	Yes
New Hampshire	No	No	No
New Jersey	Yes	No	No
New Mexico	No	No	No
New York	Yes	Yes	No
North Carolina	No	Yes	Yes
North Dakota	No	No	No
Ohio	No	Yes	No

(Table continued)

	State exit or end-of-course exams are based on 10th grade standards or higher	State requires remediation for students failing exit, end-of-course, or promotion exams	State has a non-standard diplomas or tiered diploma system
Oklahoma	No	No	No
Oregon	No	No	No
Pennsylvania	No	No	No
Rhode Island	No	No	No
South Carolina	Yes	Yes	No
South Dakota	No	No	No
Tennessee	Yes	Yes	Yes
Texas	Yes	Yes	No
Utah	No	No	No
Vermont	No	No	No
Virginia	Yes	No	Yes
Washington	No	No	No
West Virginia	No	No	No
Wisconsin	No	No	No
Wyoming	No	No	No
U.S.	Yes: 17	Yes: 18	Yes: 8



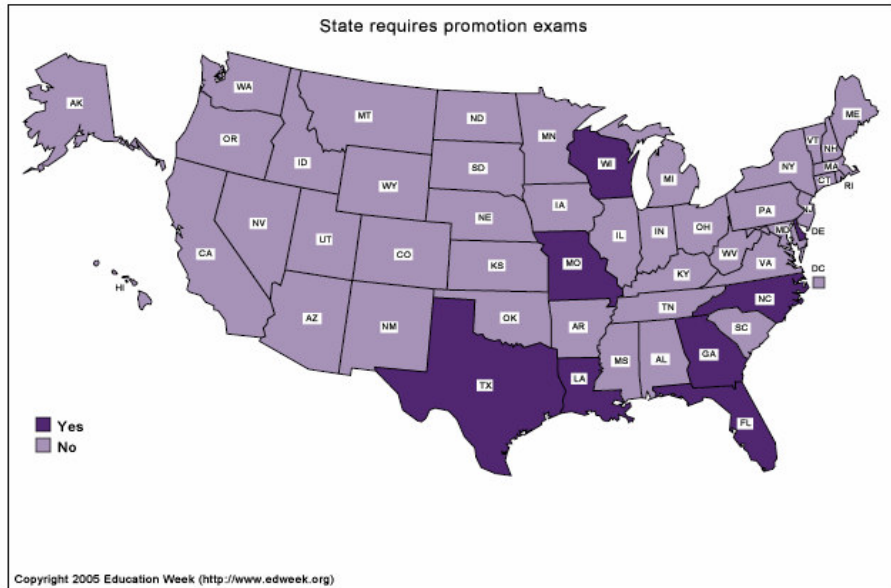


Figure A.1:  
Promotion Exams

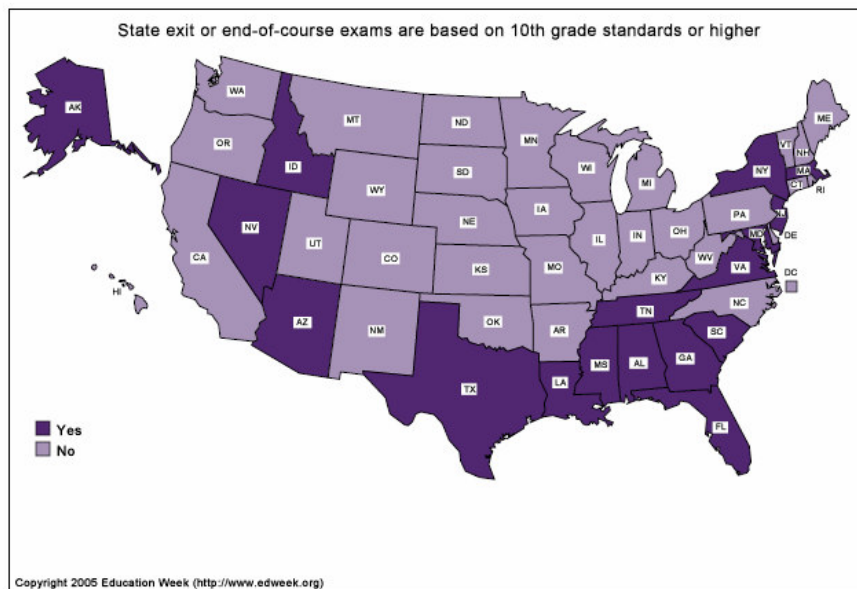


Figure A.2:  
State Exit Exams

Table A.2:

Current Participation In High Stakes Testing And Content Areas

	Elementary school English/language arts test is aligned to state standards	Middle school English/language arts test is aligned to state standards	High school English/language arts test is aligned to state standards	Elementary school mathematics test is aligned to state standards	Middle school mathematics test is aligned to state standards	High school mathematics test is aligned to state standards	Elementary school science test is aligned to state standards	Middle school science test is aligned to state standards	High school science test is aligned to state standards
Alabama	Yes	Yes	Yes	Yes	Yes	Yes	No	No	Yes
Alaska	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No
Arizona	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No
Arkansas	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No
California	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes
Colorado	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	No
Connecticut	Yes	Yes	Yes	Yes	Yes	Yes	No	No	Yes
Delaware	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District of Columbia	No	No	No	No	No	No	No	No	No
Florida	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Georgia	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Hawaii	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No
Idaho	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No
Illinois	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Indiana	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No
Iowa	No	No	No	No	No	No	No	No	No
Kansas	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Kentucky	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Louisiana	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Maine	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Maryland	Yes	Yes	Yes	Yes	Yes	Yes	No	No	Yes

(Table continued)

Massachusetts	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No
Michigan	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Minnesota	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No
Mississippi	Yes	Yes	Yes	Yes	Yes	Yes	No	No	Yes
Missouri	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No
Montana	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No
Nebraska	Yes	Yes	Yes	No	No	No	No	No	No
Nevada	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No
New Hampshire	No	No	Yes	No	No	Yes	No	No	No
New Jersey	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	No
New Mexico	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
New York	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
North Carolina	Yes	Yes	Yes	Yes	Yes	Yes	No	No	Yes
North Dakota	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No
Ohio	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Oklahoma	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Oregon	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pennsylvania	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No
Rhode Island	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No
South Carolina	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
South Dakota	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Tennessee	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Texas	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes
Utah	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Vermont	Yes	No	Yes	No	No	Yes	No	No	No
Virginia	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Washington	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
West Virginia	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Wisconsin	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Wyoming	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No
U.S.	Yes: 48	Yes: 47	Yes: 49	Yes: 46	Yes: 46	Yes: 48	Yes: 27	Yes: 26	Yes: 29

## **APPENDIX B: EFFECTS OF HIGH STAKES TESTING**

Table B.1:

Effects on Curriculum and Instruction

Provided by Langenfeld, Thurlow, and Scott (1997)

Study	Type of Test	Method	Subjects	Results
Berger & Elson (1996). What happens when MCTs are used as an accountability device: Effects on teaching autonomy, cooperation and school mission (also in Attitudes and School Climate)	Graduation exit exam	Survey. Compared responses of teachers in high stakes programs (graduation exit exam) to teachers in low stakes programs.	National Schools and Staffing Survey (SASS), Dept. of Ed. 1987. Representative national sample of teachers from low and high-stakes states.	Loss of teacher autonomy; increased clarity of mission; no effect on teacher cooperation; increased emphasis on basics.
Bergquist, Elzie & Groves (undated). Evaluation of the impact and effectiveness of recent changes in Florida's graduation and competency test standards on the educational opportunities provided handicapped students.	Increased standards and graduation requirements. Did not report on the effects of the graduation exit exam.	Program evaluation. Student records, surveys, and qualitative observations.	93 onsite visits statewide. 300 interviews from all levels of system. Reviewed more than 100 student records. Surveyed 500 knowledgeable people. Onsite visits to nine representative districts.	Students with disabilities had difficulty meeting the higher standards and incorporating nonacademic subjects into the curriculum.
Grossman, Kirst, & Schmidt-Posner (1986). On the trail of the omnibeast: Evaluating omnibus education reforms in the 1980s.	California education reforms including increased graduation requirements, increased college entrance requirements.	Examined course offerings from 1982 to 1985.	Sampled course offerings in 200 school districts statewide.	Increased offerings in academic areas, especially math, science, and advanced placement; decreased offerings in industrial arts, home economics and business ed.
Herman & Golan (undated). Effects of standardized testing on teachers and learning - another look.	Standardized testing in nine states.	Survey	Teachers chosen in matched pairs from demographically similar schools that had shown substantial improvement on standardized tests and those which had not. Total of 24 pairs.	Substantial time and effort devoted to testing; pressure from schools to improve scores; modification of curriculum; greater impact in lowest SES schools. Gains not clear as to real improvement or teaching to the test.

Study	Type of Test	Method	Subjects	Results
Shepard (1990). "Inflated test score gains." Is it old norms or teaching to the test?	All	Survey of state directors of testing regarding narrowing of the curriculum and teaching to the test.	State testing directors in the 46 states that had testing programs.	Found instances of teaching to the test. Could not quantify how teaching to the test had affected test scores. Recommended new tests every year.
Shepard & Dougherty (1991). Effects of high stakes testing on instruction.	Standardized test (two states, unspecified). Third, fifth, and sixth grade testing programs.	Survey	Teachers in districts with high stakes testing. 360 teachers responded (42% return rate).	Greater emphasis on basic skills; neglect of non-tested material; inordinate time on test preparation throughout year; clearer instruction goals.

Table B.2:

Effects on Student Learning

Study	Type of Test	Method	Subjects	Results
Allington & McGill-Franzen (1992a). Improving school effectiveness or inflating high-stakes test performances?	Third grade testing in reading and mathematics.	Examination of student records using students' scores by kindergarten cohort rather than grade.	Seven schools (low, moderate and high poverty levels). 11 - 70 subjects in each school. Total subjects not reported.	Data suggests that reported test score gains result from retention and referral to special education.
Allington & McGill-Franzen (1992b). Unintended effects of educational reform in New York.	Third grade testing in reading and mathematics.	Analyses of trends in special education referral and grade retention.	Seven school (low, moderate and high poverty levels). 11 - 70 subjects in each school. Total subjects not reported.	Data suggest increasing trends in special education referral and grade retention during increased stakes.
Catterall (1987). Standards and school dropouts: A national study of the minimum competency test.	Graduation exit exam.	Interviews with educators and students in four states with graduation exit exams.	736 students. Six state test directors. 13 district test coordinators. 18 school principals. 21 school counselors.	Professionals did not believe that the exit tests had affected dropout rates. Students, while supportive of testing in general reported more effects on dropout rates. Found association between competency test failure and reduced beliefs that the student would finish school.
Corbett & Wilson (1991). Testing reform and rebellion (also in Attitudes and School Climate). Corbett & Wilson (1990). Unintended and unwelcome: The local impact of state testing.	Graduation exit exam (Maryland). Exam to identify students in need of remediation (Pennsylvania).	Compared high-stakes state (MD) to low stakes (PA). Qualitative interviews, extended site visits, and surveys. Twelve sites visited (six from each state).	Over 250 educators.	Greater impact in high stakes state; narrowed curriculum; greater emphasis on basic skills; neglect of non-tested subjects; increased preparation time for tests; increased clarity of educational goals.

<b>Study</b>	<b>Type of Test</b>	<b>Method</b>	<b>Subjects</b>	<b>Results</b>
Mangino, Battaile, & Washington (1986). Minimum competency for graduation.	Graduation exit exam.	Analyzed data from the Texas graduation exit exam 1984 - 1985.	184 students in reading and 115 in math (total number of students receiving waivers).	Problem areas found, including taking the test many times (up to 4), waivers, and special education exemptions.
Morris (1991). Structural pattern and change in grade retention rates: An aggregate analysis of data from a large urban school district, 1982 - 1989.	Test used initially to identify weak students, later used to identify weak programs (from low to high stakes).	Analyzed data from Florida's minimum competency testing programs.	All students taking the exam.	Increased retention rates due to school restructuring and increased standards.
Potter and Wall (1992). Higher standards for grade promotion and graduation: Unintended effects of reform.	Graduation exit exam, grade promotion.	Longitudinal study of effects of SC graduation, promotion exams, 1985 - 91.	All students taking SC exams.	Students were more likely to be retained (were overage for their grade) as a result of high-stakes testing.
Griffin & Heidorn (1996). An examination of the relationship between minimum competency test performance and dropping out of high school.	Graduation exit exam.	Examined the effects of graduation exit exam on dropout rates in Florida.	Cross-sections, random sample of students in high school from 14 school districts, grades 10, 11, and 12. N=76,664 students in 75 high schools.	Students who did not pass the MCT test were not more likely to drop-out, regardless of SES, other factors. Exception: students with high GPAs who failed the exam were more likely to drop out.
Reardon (1996). Eighth grade minimum competency testing and early high school dropout patterns.	Promotion from eighth grade to ninth grade.	Examined data from National Educational Longitudinal Survey (NELS).	Focused on MCT in eighth grade. Nationwide representative sample.	Increased dropout rates in programs with eighth grade MCT.
Walstad (1984). Analyzing minimal competency test performance.	Test used to gauge district/student performance.	Used an educational production function model based on prior research and two years of MCT data. Surveys of district administrators.	District administrators. Used data from all students taking MCT in state.	Only pretesting (practicing the test) had a significant effect on student performance. Changes in curricula and teacher training had no effect on scores.
Chin-Chance, Gronna & Jenkins (1996). Assessing special education students in a norm referenced statewide testing program.	SAT8 (Stanford Achievement Test 8). Used for district comparisons.	Developed separate norms, percentile ranks for students with disabilities.	Students with disabilities (all taking exam).	Were able to develop separate percentile ranks. While students without disabilities, difference scores showed students with disabilities improved as much or more than students without disabilities.
Safer (1980). Implications of minimum competency standards and testing for handicapped students.	Graduation exit exam.	Examined scores of students with disabilities taking graduation exit exam.	All students with disabilities taking the exam in 1977.	Students with disabilities were not likely to pass the exam.

<b>Study</b>	<b>Type of Test</b>	<b>Method</b>	<b>Subjects</b>	<b>Results</b>
Vitello, Camilli & Molenaar (1987). Performance of special education students on a minimal competency test.	Graduation exit exam.	Analyzed scores of students with disabilities taking the NJ proficiency exam.	All 4,299 students with disabilities who took exam 1986 - 87 (40% of the ninth grade handicapped population).	Students with disabilities were not likely to pass.

Table B.3:

Effects on Attitudes and School Climate

<b>Study</b>	<b>Type of Test</b>	<b>Method</b>	<b>Subjects</b>	<b>Results</b>
Berger & Elson (1996). What happens when MCTs are used as an accountability device: Effects on teaching autonomy, cooperation and school mission (also in Curriculum).	Graduation exit exam.	Survey. Compared responses of teachers in high stakes programs (graduation exit exam) to teachers in low stakes programs.	National Schools and Staffing Survey (SASS), Dept. of Ed. 1987. Representative national sample of teachers from low and high-stakes states.	Loss of teacher autonomy; increased clarity of mission; no effect on teacher cooperation; increased emphasis on basics.
Corbett & Wilson (1991). Testing reform and rebellion (also in Curriculum).	Graduation exit exam (Maryland). Exam to identify students in need of remediation (Pennsylvania).	Compared high-stakes state (MD) to low stakes (PA). Qualitative interviews, extended site visits, and surveys. Twelve sites visited (six from each state).	Over 250 educators.	Greater impact in high stakes state; narrowed curriculum; greater emphasis on basic skills; neglect of non-tested subjects; increased preparation time for tests; increased clarity of educational goals.
Rottenberg & Smith (1990). Unintended effects of external testing in elementary schools (also in Student Learning).	ITBS used for evaluation of principals and making curriculum decisions. Media coverage of test scores.	Interviews with teachers. Dual case study design.	19 teachers in two school districts. Observation of four local teachers.	Negative effects on the attitudes of both teachers and students. Reduced time for ordinary instruction.
Rodgers, Parededs, & Mangino (1991). High Stakes minimum skills tests: Is their use increasing achievement?	Graduation exit exam.	Used multivariate analysis of variance (MANOVA) on the year of the test, GPA, year X GPA interaction, and GPA squared. Expected basic skills to increase while higher order skills decreased.	12,404 11th grade students in the Austin Independent School District. Special education and LEP students were not included in the study.	Basic skills increased while higher order skills remained the same.



## **APPENDIX C: SURVEY INSTRUMENT**

Directions: In this questionnaire, the term “test” refers to the LEAP and iLEAP tests. Please respond to the following questions by either circling or filling in your responses.

1. School name: \_\_\_\_\_
  
2. Including this year, how many years have you been teaching? \_\_\_\_\_  
(1) 0-3 (2) 4-6 (3) 7-9 (4) 10-12 (5) 13 or more years
  
3. What grade levels are you teaching this year? (Circle all that apply)  
1 2 3 4 5 \_\_\_\_\_
  
4. What subjects do you teach? (Circle all that apply)  
English/Language Arts (1)    Mathematics (2)    Social Studies (2)    Science (4)  
Elementary (5 – all subjects)  
Other, please specify \_\_\_\_\_
  
5. Do you teach more than one class a day? Yes (1) No (2) [Teachers that stay with the same students all day have one class] \_\_\_\_\_
  
6. Despite LEAP/iLEAP testing, I have control over my classroom program.  
Strongly Disagree                      Neutral                      Strongly Agree  
1                      2                      3                      4                      5
  
7. Testing has affected my commitment to the education profession.  
Strongly Disagree                      Neutral                      Strongly Agree  
1                      2                      3                      4                      5
  
8. Testing has affected my satisfaction with my work.  
Strongly Disagree                      Neutral                      Strongly Agree  
1                      2                      3                      4                      5
  
9. Due to testing I am more likely to choose another profession.  
Strongly Disagree                      Neutral                      Strongly Agree  
1                      2                      3                      4                      5
  
10. I perceive I am being judged solely on how my students perform on the LEAP/iLEAP test.  
Strongly Disagree                      Neutral                      Strongly Agree  
1                      2                      3                      4                      5
  
11. LEAP/iLEAP testing has changed content of instruction in your classroom.  
Strongly Disagree                      Neutral                      Strongly Agree  
1                      2                      3                      4                      5

12. I make sure the objectives of the test are covered in my instruction.

Strongly Disagree		Neutral		Strongly Agree
1	2	3	4	5

13. Our school is more interested in increasing test scores than improving overall student learning.

Strongly Disagree		Neutral		Strongly Agree
1	2	3	4	5

14. Testing has affected my ability to meet students' individual needs.

Strongly Disagree		Neutral		Strongly Agree
1	2	3	4	5

15. I have made changes in my classroom instruction in the last few years in response to LEAP or iLEAP testing.

Strongly Disagree		Neutral		Strongly Agree
1	2	3	4	5

16. I am pressured to cover all the required curriculum that is covered by LEAP/iLEAP testing.

Strongly Disagree		Neutral		Strongly Agree
1	2	3	4	5

17. I perceive pressure to teach test preparation.

Strongly Disagree		Neutral		Strongly Agree
1	2	3	4	5

18. My school considers test scores when evaluating teachers.

Strongly Disagree		Neutral		Strongly Agree
1	2	3	4	5

19. To what extent do you perceive pressure from the following groups to improve your students' LEAP/iLEAP test scores:

	Almost No Pressure		Moderate		Great Pressure
a. Principal	1	2	3	4	5
b. Other School Administrators	1	2	3	4	5
c. Other Teachers	1	2	3	4	5
d. Central Office	1	2	3	4	5
e. Parents	1	2	3	4	5
f. Newspaper/Media	1	2	3	4	5

20. LEAP/ iLEAP testing is helping schools to improve.

Strongly Disagree		Neutral		Strongly Agree
1	2	3	4	5

21. Teachers who complain about testing are usually poor teachers who do not wish to be accountable as professionals.

Strongly Disagree

Neutral

Strongly Agree

1

2

3

4

5

**22. What are two positive and two negative consequences of LEAP/iLEAP testing?**

Thank you for taking the time to complete this survey. If you would like a copy of results, please provide me with your address and we will send you a copy.

## **APPENDIX D: INTERVIEW PROTOCOL**

**Opening Statement:** the purpose of this interview is to obtain information than can enhance our understanding of the effects of LEAP21 testing on school improvement. As an educator who has had experience in a high stakes testing environment, you are in a unique position to describe these effects and their impacts on the school setting. The information that is gained from these interviews will be used in papers and presentations related to high stakes testing. No real names will be used in this study, as noted on the consent form. If you would like a copy of the paper, I would be happy to provide you with one. As we go through the interview, if you have any questions about why I am asking you something or if you need further clarification, please feel free to ask. Do you have any questions before we begin?

**Questions:**

1. How does test preparation (teaching the test) affect your instructional planning, learning strategies, and curriculum content?
2. To what extent does test preparation (teaching the test) affect your instructional planning, learning strategies, and curriculum content?
3. How much time do you perceive that students spend on test preparation and how does that amount of time compare to the time spent on instruction?
4. What effect does testing have on an educators' sense of professionalism and pride in their work?
5. How does high stakes testing affect motivation in general?
6. From where do you perceive pressure to increase scores?
7. Has this increased attention been positive or negative toward the students' academic achievement?

8. In what direction do you believe education is heading into since the implementation of LEAP21 testing?

**APPENDIX E: PERMISSION LETTER**



To: XXX XXXX (Principal)  
1234 XXX Blvd.  
Marrero, LA 70072

From: David C. Charles, Ph.D. Student  
Louisiana State University  
Harvey, LA 70058

I have received permission to gather data for a research project as a part of my Ph.D. program at Louisiana State University. Likewise, I have attached a narrative discussion of my dissertation, which provides some information about the study I would like to conduct in your school and the potential benefits the results might contribute to the improvement of education.

Please submit the provided surveys to your teachers on **August 8<sup>th</sup>**. The surveys are not intrusive and take on average between 4 and 5 minutes to complete. After a lead teacher collects the surveys, please place all completed and blank copies in the manila folder provided and pony it to **Attn: David Charles at Woodmere Elementary**.

At this point, I have completed all of my course work in my Ph.D. program of studies in Educational Leadership, Research and Counseling at Louisiana State University. I have successfully defended my research proposal at LSU.

The following is a brief overview of the study:

**Title of the Research Study:** Teachers' Perspectives on the Unintended Consequences of High Stakes Testing

**Research Director:** Student Director: David Charles

**Purpose of the Study:** This study was organized as an exploratory investigation of the effects of high stakes testing on the area of morale, pressure, and classroom practices as perceived through the experiences of teachers.

**Procedures to be Used:** The teachers will complete a short survey. Shortly after, two teachers from each school will meet with the researcher for a short interview at their school.

**Potential Risks to Participants:** There is no apparent risk to the participants involved in this study.

**Potential Benefits of the Study:** By identifying the impact of accountability on instruction and teachers, schools administrators and parents/guardians can develop strategies to increase the student success rate on the criterion-referenced test.

**Protection of the identity and privacy of the participants:** The teachers are instructed to answer only the questions on the instrument and not add any additional markings. Other than the survey questions, only general demographic information will be asked. Teachers will be asked to complete instrument and return it to a lead teacher, which will be sealed and given to the investigator. Once returned to the investigator, the instrument will be sorted by school for analyses.

Thank you in advance for your support. I can be reached by phone at home at 985-725-0449 or work 504-371-0476 or by e-mail at [David.Charles@jppss.k12.la.us](mailto:David.Charles@jppss.k12.la.us) or [dccharles@yahoo.com](mailto:dccharles@yahoo.com).

Sincerely,  
David C. Charles

---

**APPENDIX F: INTERVIEW PERMISSION LETTER**

To: XXX XXXX (Principal or Participant)

From: David C. Charles, Ph.D. Student  
Louisiana State University  
Harvey, LA 70058

I would like to thank you for the permission to gather data for a research project as a part of my Ph.D. program at Louisiana State University. Likewise, I have attached a narrative discussion of my dissertation, which provides some information about the study I would like to conduct in your school and the potential benefits the results might contribute to the improvement of education.

The purpose of this interview is to obtain information that can enhance our understanding of the effects of LEAP21 testing on school improvement. As an educator who has had experience in a high stakes testing environment, you are in a unique position to describe these effects and their impacts on the school setting. The information that is gained from these interviews will be used in papers and presentations related to high stakes testing. No real names will be used in this study, as noted on the consent form. If you would like a copy of the paper, I would be happy to provide you with one. As we go through the interview, if you have any questions about why I am asking you something or if you need further clarification, please feel free to ask.

At this point, I have completed all of my course work in my Ph.D. program of studies in Educational Leadership, Research and Counseling at Louisiana State University. I have successfully defended my research proposal at LSU.

The following is a brief overview of the study:

**Title of the Research Study:** Teachers' Perspectives on the Unintended Consequences of High Stakes Testing

**Research Director:** Student Director: David Charles

**Purpose of the Study:** This study was organized as an exploratory investigation of the effects of high stakes testing on the area of morale, pressure, and classroom practices as perceived through the experiences of teachers.

**Procedures to be Used:** The teachers will be interviewed at \_\_\_\_\_ a.m./p.m. on \_\_\_\_\_ at your school. Please provide a place for the interviews that will be comfortable for the teachers and that will provide the least amount of interruptions. The interviews will last between 20 – 30 minutes.

**Potential Risks to Participants:** There is no apparent risk to the participants involved in this study.

**Potential Benefits of the Study:** By identifying the impact of accountability on instruction and teachers, schools administrators and parents/guardians can develop strategies to increase the student success rate on the criterion-referenced test.

**Protection of the identity and privacy of the participants:** The teachers are instructed to answer only the questions they are comfortable answering. Other than the interview questions, only general demographic information will be asked.

Thank you in advance for your support. I can be reached by phone at home at 985-725-

0449 or work 504-371-0476 or by e-mail at David.Charles@jppss.k12.la.us or dccharles@yahoo.com.

Sincerely,

David C. Charles

**APPENDIX G: INSTITUTIONAL REVIEW BOARD APPROVAL**

**Application for Exemption from Institutional Oversight**

Unless qualified as meeting the specific criteria for exemption from Institutional Review Board (IRB) oversight, ALL LSU research/projects using living humans as subjects, or samples or data obtained from humans, directly or indirectly, with or without their consent, must be approved or exempted in advance by the LSU IRB. This Form helps the PI determine if a project may be exempted, and is used to request an exemption.



Institutional Review Board  
 Dr. Robert Mathews, Chair  
 203 B-1 David Boyd Hall  
 Baton Rouge, LA 70803  
 P: 225.578.8692  
 F: 225.578.6792  
 irb@lsu.edu | lsu.edu/irb

- Applicant, Please fill out the application in its entirety and include the completed application as well as parts A-E, listed below, when submitting to the IRB. Once the application is completed, please submit two copies of the completed application to the IRB Office or to a member of the Human Subjects Screening Committee. Members of this committee can be found at [http://appl003.lsu.edu/osp/osp.nsf/\\$Content/Humans+Subject+Committee?OpenDocument](http://appl003.lsu.edu/osp/osp.nsf/$Content/Humans+Subject+Committee?OpenDocument)
- A Complete Application Includes All of the Following:
  - (A) Two copies of this completed form and two copies of parts B thru E.
  - (B) A brief project description (adequate to evaluate risks to subjects and to explain your responses to Parts 1 & 2)
  - (C) Copies of all instruments to be used.
    - If this proposal is part of a grant proposal, include a copy of the proposal and all recruitment material.
  - (D) The consent form that you will use in the study (see part 3 for more information.)
  - (E) Certificate of Completion of Human Subjects Protection Training for all personnel involved in the project, including students who are involved with testing or handling data, unless already on file with the IRB.  
 Training link: ( <http://cme.cancer.gov/clinicaltrials/learning/humanparticipant-protections.asp>.)

1) Principal Investigator: David Charles Rank: Ph.D. Candidate Student\*? Y/N

Dept.: ELRC Ph: 985-725-0449 E-mail: dccharles@yahoo.com

2) Co Investigator(s): please include department, rank and e-mail for each  
 If student, please identify and name supervising professor in this space  
None

3) Project Title: Teachers' Perspectives on the Unintended Consequences of High Stakes Testing

4) LSU Proposal?(yes or no) No If Yes, LSU Proposal Number \_\_\_\_\_  
 Also, if YES, either  This application completely matches the scope of work in the grant  
 OR  More IRB Applications will be filed later

5) Subject pool (e.g. Psychology Students) Elementary School Teachers  
 •Circle any "vulnerable populations" to be used. (children <18; the mentally impaired, pregnant women, the aged, other). Projects with incarcerated persons cannot be exempted.

6) PI Signature [Signature] \*\* Date 9/18/07 (no per signatures)  
 \*\*I certify my responses are accurate and complete. If the project scope or design is later changed I will resubmit for review. I will obtain written approval from the Authorized Representative of all non-LSU institutions in which the study is conducted. I also understand that it is my responsibility to maintain copies of all consent forms at LSU for three years after completion of the study. If I leave LSU before that time the consent forms should be preserved in the Departmental Office.

\*\*\*Effective August 1, 2007, all Exemptions will expire three years from date of approval, unless a continuation report, found on our website, is filed prior to expiration date\*\*\*

Study Exempted By:  
 Dr. Robert C. Mathews, Chairman  
 Institutional Review Board  
 Louisiana State University  
 203 B-1 David Boyd Hall  
 225-578-8692 | www.lsu.edu/irb  
 Exemption Expires: 09-18-2010

IRB# 3715 LSU Proposal# \_\_\_\_\_  
 Complete Application  
 Human Subjects Training

Screening Committee Action: Exempted  Not Exempted \_\_\_\_\_ Category/Paragraph 1  
 Reviewer Mathews Signature [Signature] Date 9/20/07

**APPENDIX H: JEFFERSON PARISH PUBLIC SCHOOL SYSTEM APPROVAL**



**DIANE M. ROUSSEL, Ph.D.**  
SUPERINTENDENT

**INSTRUCTIONAL SUPPORT SERVICES  
JEFFERSON PARISH PUBLIC SCHOOL SYSTEM**

4600 RIVER ROAD  
MARRERO, LOUISIANA 70072  
(504) 349-7964  
FAX: (504) 349-7797  
www.jpss.k12.la.us

TITLE II  
FOR  
TITLE V  
INSTRUCTION

**MARIAN BERNARD**  
DIRECTOR OF CURRICULUM AND INSTRUCTION  
AND PROFESSIONAL LEARNING

September 20, 2007

David Charles  
223 Ormond Meadows Drive  
Destrehan, LA 70047

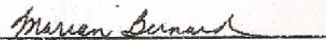
Dear Mr. Charles,

This is to notify you that your request to conduct research in Jefferson Parish has been approved with the following qualifications:

1. You must obtain permission to conduct the study from the principal of the school(s) involved. You should present this letter as your official district approval. The principal may request the research packet. You must notify this office of the school(s) selected for participation prior to the study, 349-7964; fax 349-7797; or e-mail [marian.bernard@jpss.k12.la.us](mailto:marian.bernard@jpss.k12.la.us).
2. You need to be certain all participants understand their involvement is voluntary and that the confidentiality of reported data will be maintained.
3. Send a copy of the completed report or paper with the findings of your research project. Your findings will add to our body of knowledge.

If you have any questions about your application, or if I can be of assistance, please contact me at 349-7964.

Approved:

  
Marian Bernard  
Director of Curriculum and Instruction &  
Professional Learning

TOTAL P.02

## VITA

David Christopher Charles was born in 1964 in New Orleans, Louisiana. He graduated from Louisiana State University with a bachelor's degree in social studies/ secondary education in 1987.

David was employed as a teacher for 8 years. He was named "Outstanding Young Educator of the Year" in 1995. He served as a teacher at Destrehan High School, Marrero Middle, and East Jefferson High School. He was the originator of many programs and extracurricular activities. He also coached winning football and basketball teams.

While he served as a teacher at East Jefferson High School, he was promoted to Dean of Students. He has since moved on to Vice-Principal of Roosevelt Middle, Woodmere Elementary, and George Cox Elementary and Principal of Henry Ford Middle School. In the last 9 years, he has served as an administrator at all levels at elementary, middle, and high schools in Jefferson Parish.

He received a master's degree in educational administration and supervision from the University of New Orleans in 1997. David received a Master +30 degree from the State while attending Louisiana State University in Baton Rouge, Louisiana.

David has been married for 14 years to wife, Colleen and has two sons: Nicky, 11 and Christopher, 5. He has been employed by the Jefferson Parish Public School System for 17 years.

David has been in the United States Army Reserves for 18 years as both an enlisted soldier and an officer. He served with distinction overseas in the Gulf war and currently holds the rank of Major. The military has enrolled him in numerous educational courses concerning logistics and leadership. He has been the commander of the 215<sup>th</sup> MPAD (Mob) for the last four years.



David currently lives in Destrehan, Louisiana. He plans to spend more time with his family upon completion of this dissertation.