UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

TEACHING WITH THE DATA IN MIND: USING OKLAHOMA ACHIEVEMENT TEST RESULTS TO CLOSE THE GAP FOR NO CHILD LEFT BEHIND

A Dissertation

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

degree of

Doctor of Education

Rebecca York
Norman, Oklahoma
2004

Downloaded from SHAREOK. Persistent Handle URL: https://hdl.handle.net/11244/694

UMI Number: 3120033

Copyright 2004 by
York, Rebecca

All rights reserved.

INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleed-through, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.



UMI Microform 3120033
Copyright 2004 by ProQuest Information and Learning Company.
All rights reserved. This microform edition is protected against unauthorized copying under Title 17, United States Code.

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

© Copyright by Rebecca York, 2004
All Rights Reserved

TEACHING WITH THE DATA IN MIND
USING OKLAHOMA ACHIEVEMENT TEST RESULTS TO
CLOSE THE GAP FOR NO CHILD LEFT BEHIND

A Dissertation APPROVED FOR THE
DEPARTMENT OF EDUCATIONAL LEADERSHIP AND POLICY STUDIES

BY

Dr. Gregg Garn, Chair

Dr. Jeff Maiden

Dr. Courtney Vaughn

Dr. Frank McQuarrie

Dr. Gaetane Jean-Marie

ACKNOWLEDGEMENTS

With deep appreciation and humility, I would like to thank several people who made my doctoral program a pleasant and successful journey. First, Dr. Gregg Garn agreed to be my committee chair, offering valuable, practical guidance where needed and giving willingly and patiently of his time and infinite wisdom to answer my many questions. From the beginning, Dr. Jeffrey Maiden believed in me, giving me the encouragement I needed to pursue the degree and offering his assistance numerous times along the way, including being a member of my committee. As my professor and then a committee member, Dr. Courtney Vaughn correctly advised me to consider my strengths and undertake this qualitative study. Always with helpful suggestions, as an outstanding teacher in the program and a committee member, Dr. Frank McQuarrie challenged me at every stage of the process to reach a little deeper and achieve more than I thought possible. With a high regard for the public school teacher, Dr. Gaetane Jean-Marie, a vital committee member, supported my efforts to pursue the effective classroom teacher as a "wise expert" and true specialist in the profession.

In addition, I am truly grateful to the six outstanding teachers who gave their precious time and participated in the detailed interviews. They truly were "wise experts" and rare connoisseurs of the teaching/learning process; I now consider them lifelong friends as well as colleagues. Their thoughts, their trust, and their time not only made this research possible, but very rewarding.

DEDICATION

This manuscript is dedicated

To my late Cherokee mother and grandmothers who continued to guide me

In the spirit, challenging me to climb "tall trees," and

To my husband and family, without whose support

I could not have climbed this tallest of all "trees."

TABLE OF CONTENTS

CHAPTER I INTRODUCTION Page

Research Problem Introduction 1

Overview of National Educational Achievement Tests 2
Goals 2000 4
No Child Left Behind Act of 2001 5
Widely Used Achievement Tests 6

Overview of Oklahoma Achievement Tests 7
PASS Standards 7
Oklahoma State Assessment Policies 8
Choosing Assessments That Measure Standards 8
No Child Left Behind Requirements for Oklahoma 10

Problems using Oklahoma Assessment Data 12
History of Oklahoma Test Vendors 13
Test Results Decide Low-Performing Schools 16

Overview of Oklahoma's Academic Performance Index 17
Oklahoma Plans for Future Achievement Tests 19

Significance of the Study 20

Research Question 21

Scope and Participants of the Study 21

Statement of the Problem 22

Limitations of the Study 22

Definition of Terms 23

CHAPTER II LITERATURE REVIEW

The Arguments For and Against Standardized Testing 26
Advocates for Standardized Testing 27

Critics of Standardized Testing..... 30
Testing Scandals..... 32
Problems with Standardized Testing..... 33
Educational Accountability..... 34
Could Test Policy Change Practice..... 35

Current Rewards/Sanctions Provided by NCLB.....36
Adequate Yearly Progress..... 36
Highly Qualified Teachers..... 40
High Stakes Decisions for Oklahoma..... 40
Oklahoma Test Contracts and Vendors..... 41

Moving Beyond Current Student Assessments 42
Becoming Data-Driven..... 43
Assessments That Improve Learning 43
Purpose of Assessments..... 45
Use of Previous Oklahoma Test Data.....47

Norm-Referenced versus the Criterion-Referenced Tests.....47

Key Components of Oklahoma Academic Performance Index..... 52
API Formulas and Growth Targets.....54
Test Results Used for the Oklahoma API.....56
Changes in Assessment Process..... 58

How the Oklahoma API Compares to Other States..... 60
Vermont's Accountability System.....62
Nebraska's Accountability System.....64
The Texas Accountability System.....66
Other States' Accountability Systems..... 70

CHAPTER III METHODOLOGY AND DESIGN

Introduction.....74
Research Question..... 74

Case Studies.....75
Background for the Case Study..... 77

Data Sources..... 78
Survey..... 78
Interviews..... 80
Observations..... 85
Documents..... 87

Data Analysis..... 87
Transferability and Replicability..... 89
Triangulation..... 90

The Researcher..... 93

Community Context..... 94

Communicating the Case Study.....95

CHAPTER IV RESULTS OF THE RESEARCH

Introduction.....97

Information Obtained from Initial Survey.....97

Summary of the Surveys.....98

Interviews with Six Key Elementary Teachers..... 101
Introduction to Six Key Teachers..... 102

Results of Interviews with Key Participants..... 109
Theme 1: Oklahoma State Standards.....110
Theme 2: Oklahoma Assessments and Blueprints.....115
Theme 3: Teachers' Use of Educational Data..... 123
Theme 4: Proven Practices/Interventions.....129
Theme 5: The Issue of Time.....134
Theme 6: Discipline and Parental Involvement.....139
Theme 7: High Expectations for Students.....144

Other Issues or Non-Issues that Effect Student Achievement.....148

CHAPTER V CONCLUSIONS/RECOMMENDATIONS

Introduction.....152

The Initial Survey.....154

Theme 1: Conclusions about Oklahoma State PASS Standards.....155
Recommendations.....157

Theme 2: Conclusions about State Assessments and Blueprints.....158
Recommendations.....162

Theme 3: Conclusions on Evaluation/Use of Educational Data.....163
Recommendations.....166

Theme 4: Conclusions on Proven Practices and Interventions.....166
Recommendations.....167

Theme 5: Conclusions on Issue of Time for Using Data.....168
Recommendations.....171

Theme 6: Conclusions on Discipline and Parent Involvement.....171
Recommendations.....173

Theme 7: Conclusions on High Expectations for Students.....173
Recommendations.....175

Topic 8: Conclusions on Other Issues/Non-Issues.....175
Recommendations.....177

Implications from Other States for Oklahoma School Leaders 177

Final Recommendations.....180
Recommendations for Further Study.....181

REFERENCES.....183

APPENDIX.....189

LIST OF TABLES

Table		Page
1.	1994-2003 History of Oklahoma State-Mandated Tests	15
2.	NCLB School Improvement Steps by Year	38
3.	Difference in NRT and CRT Scores	49
4.	2001 Academic Performance Index for Oklahoma	53
5.	Examples of Some API Formulas used in the 2001 API	55
6.	How FairTest Scored the States	62
7.	Years and Grade Levels of Teachers Interviewed	102
8.	Matrix of Findings on Oklahoma State Standards	114
9.	Matrix of Findings on Oklahoma State Assessments/Blueprints	122
10.	Matrix of Findings for Teachers' Use of Educational Data	128
11.	Matrix of Findings on Teachers' Proven Practices/Interventions	133
12.	Matrix of Findings for the Issue of "Time" for Teachers	139
13	Matrix of Findings for Discipline and Parent Involvement	144
14.	Matrix of Findings for High Expectations of Students	148
15.	Proven Practices of Highly Effective Teachers	167

LIST OF APPENDICES

Appendix	Page
A. Survey Instrument and Results for all District Third and Fifth Grade Teachers with Survey Graphs on Each Question	189-99
B. PASS Standards Documented	200
C. Oklahoma PASS Blueprint for Grade 5 Reading	201
D. Group Report for Third Grade Reading/Math/Language	202
E. Diagnostic Computer Report for Multiple Students	203
F. Multiple Lesson Report from Computer for One Student	204
G. Teacher Made Assessment/Benchmark	205
H. Computer Generated Item Analysis for Reading	206
I. Teacher-Made Assessment by Skill	207
J. Chart of Student Progress by Letter Grade	208
K. Computer Generated Lesson	209
L. Calendar of After-School Duties and Meetings	210
M. Teacher Documentation of Discipline Events	211
N. Documentation of Parent Involvement	212-13
O. Teacher Document of High Expectations	214
P. Research Questions Used in Interviews	215

ABSTRACT

This qualitative study investigated successful practices of "highly effective" elementary teachers as they used Oklahoma state achievement test data to close the gap for all student subgroups in fulfilling the mandates of the No Child Left Behind Act of 2001 (NCLB). Data from an initial survey of 30 "highly effective" third and fifth grade teachers in one Oklahoma district, in depth interviews with six of the 30 survey participants, observation field notes, and various teacher artifacts were collected and analyzed. Using computer assisted qualitative analysis software NVivo, and steps outlined by Merriam (1998), seven themes emerged from the data. The six "highly effective" teachers shared proven practices in using state and classroom data that included: (1) Oklahoma State Standards; (2) Oklahoma Assessments and Blueprints; (3) Evaluation and Use of Data; (4) Classroom Interventions and Benchmarks; (5) Time for Planning; (6) Discipline and Parental Involvement; and (7) High Expectations for Students.

INTRODUCTION

Chapter I

Research Problem Introduction

For Oklahoma, state achievement tests became "high stakes" in 2001 when state legislators included the reading and math results from those tests in the Oklahoma Academic Performance Index (API). While some states created new tests for a new system, Oklahoma used existing tests to establish the API baseline and meet the national mandates for the No Child Left Behind Act of 2001 (NCLB). The reading and math results from the state tests constituted 80-90% of the API that ranked Oklahoma school districts. Scores from 0-1500 ranked districts and schools in the state and resulted in either rewards or sanctions.

Oklahoma policy makers had implemented a "hodgepodge" (Gehring, 2002) of different tests, including both norm-referenced and criterion-referenced in the last ten years, 1994-2004. The Oklahoma State Department of Education planned to drop the norm-referenced test at the third grade level in 2005, and grades three through eight would have only criterion-referenced tests; however, Oklahoma teachers were forced to use existing data to meet the federal mandates of Adequate Yearly Progress in the No Child Left Behind legislation. While the state assessments remained in a state of constant flux, classroom teachers were asked to use those results with (1) the types of tests changing from year to year and (2) the tests being written by different vendors in different years. Oklahoma

teachers struggled to find ways to use the state achievement data effectively, and they relied on other methods and data to prepare their students in the teaching/learning process. If the state achievement tests provided constructive assessments for use by Oklahoma teachers that would also be equitable in measuring the actual progress of Oklahoma students in meeting Adequate Yearly Progress under NCLB, all teachers needed to be trained in using that data.

Overview of National Educational Achievement Tests

In public education, the term "assessment" had become a common expression to describe the various methods of evaluating students' skills and knowledge or their actual achievement. During the twentieth century, schools used a variety of achievement tests to accomplish what had become termed "accountability" for students, educators, and administrators in the task of educating children, with various states developing assessments they believed would best measure learning. The types of achievement tests usually administered to public school students measured skills and knowledge that had already been acquired. Most of those tests were multiple-choice while a few newer exams, such as the Balanced Assessments from Harvard (2004), had included open-ended or structured response questions. These achievement tests were not aptitude tests, such as the Scholastic Aptitude Tests (SAT) and Graduate Record Examinations (GRE) that predicted how well a student would perform for future classes. The achievement tests varied from state to state, and while some

state assessments had achieved some measure of success in accountability, others had not.

Achievement tests were created and used early in the twentieth century. The College Entrance Examination Board, known as the College Board (2003) began to use assessments for those who applied to colleges with the SAT first administered in June 1926. High schools and colleges all across the nation depended on the SAT for predictability of future success, and many college scholarships were awarded on those results. In 2000-2001, more than 2.1 million students took the SAT exams. In addition, the Educational Testing Service was founded in 1947 by various higher education councils to form a single national organization devoted exclusively to testing and research. However, the roots for standardized assessment had been reported at even earlier dates.

James Popham (2002) traced the ancestry of today's standardized achievement tests all the way back to World War I. The United States Army developed and implemented a testing program to identify men who would be good officer candidates. The Army Alpha intelligence test was administered to two million recruits. Popham reported that the test "performed its measurement mission with striking success" (p. 19). Because it worked so well, it became the model used for all subsequent educational tests in the United States, for both achievement and aptitude assessments.

For public schools, standardized tests became more widely used between the 1920's and 1940's period, gaining in popularity every year. As education converted from the primary-through-eighth-grade, one-room schoolhouses to larger institutions or "mass factory" models, assessment practices evolved that would show students' whole group progress. By the 1950's and 1960's, many schools, administrators, and boards of education, as well as states, used district-wide standardized testing as a way of showing progress or improvement and used their authority to pressure schools to perform well on the tests. Educational testing continued to grow into a giant industry, according to PBS Frontline (2003), with the United States Congress setting aside $400 million for states to implement third through eighth grade testing for the NCLB mandates.

Goals 2000

In the 1980's, Ronald Reagan introduced *Goals 2000*, a plan for students to demonstrate high levels of competency in core subject areas, particularly in reading and math. By the 1990's, President Bill Clinton and his secretary of education, Richard Riley, sought to align federal programs with state reforms. Riley (2002) described a "bitter battle" that ensued with critics over whether or not the United States Department of Education would survive. At the heart of this debate was a national assessment system that would, in effect, establishing a national curriculum. Wanting to end the debate over the role of the federal government in state education, Riley and the education cabinet created a "new

policy framework" with specific proposals "anchored in the standards and assessment movement" (p. 701). Those proposals created a new type of accountability for any federal monies, with Riley and Clinton asking for a voluntary national testing program.

No Child Left Behind Act of 2001

When President George W. Bush took office, he and his education secretary, Dr. Rodney Paige, introduced the No Child Left Behind Act (NCLB), a plan that reauthorized the Elementary and Secondary Act of 1965 (EASA), legislation that outlined the federal government's role in education and tied federal monies to each state's implementation of high standards and systems that test third through eighth grade students in reading and math. Paige (2002) said, "We must hold educators accountable...no middle ground or excuses" (p. 711). Each state was to implement clear standards for student achievement with annual assessments that would measure progress against those standards. Paige (2002) contended the blame game had gone on long enough, and parents and teachers must solve the "mystery" of who was "failing children" (p. 711). He believed that assessments would give educators evidence, "class by class, child by child," (p. 711) and the data would allow parents some choice in their children's education.

Under the NCLB plan, a school could be identified for "improvement or corrective action." Then, the district would have to give parents options, such as selecting another public school or tutoring services. Paige said, "There is no more

powerful force for change than parents armed with information and options. The No Child Left Behind law provides both" (p. 711). This law essentially changed the culture of education across the nation by creating the same types of reform that, according to Paige, had already shown results in Texas.

To provide some validity to state assessments, the National Assessment of Educational Progress (NAEP) (2003) developed a framework to random sample groups of students in each participating state. Individual student or school reports were not provided, but states received feedback about how their students had performed as compared to students across the nation. Established in 1969, NAEP was commonly identified as the Nation's Report Card. These tests contained constructed-response and problem-solving questions that required the use of calculators and other materials. This assessment process was very close to representing a national test in what was considered core curriculum.

Widely Used Achievement Tests

With the advances, current achievement tests were still patterned somewhat after the Army Alpha and the College Board's original SAT. The main difference was that they showed acquired knowledge rather than aptitude, and they took many different forms. The most widely used standardized tests in the United States were the TerraNova published by CTB McGraw-Hill, the Stanford Achievement Test (SAT9 and SAT10) by Harcourt Educational Measurement, the Metropolitan Achievement Test (MAT-8 and 9) also published by Harcourt, and

the Iowa Test of Basic Skills (ITBS) by Riverside Publishing. They shared one commonality: they measured acquired knowledge. These companies standardized their tests, which meant that they had been created and administered under a rigorous set of standards that everyone from the test maker to the test giver was required to follow.

Overview of Oklahoma Achievement Tests

Oklahoma followed the nation's shift in the 1993-94 school year, by developing and implementing the Priority Academic Student Skills (PASS) standards. The state standards were clearly outlined with a minimum criteria of skills and objectives for core subjects from kindergarten through twelfth grades. These standards were defined as competencies with expected levels of performance at various developmental stages. For the PASS standards, the state implemented a system for the documents to be updated every three years until 2003, when subject updates were placed on a six-year rotating cycle. That rotation matched both the state and local textbook adoption cycle and districts' adoptions of Comprehensive Local Education Plans.

PASS Standards

Generally, teachers across the state welcomed the PASS standards. They felt they were required to teach what was in the textbooks, and textbook publishers had decided Oklahoma's curriculum for far too long. Classroom teachers themselves were asked to implement and frequently revise the standards,

providing valuable input into the state's curriculum, and educators heard few complaints about the standards themselves. However, at the same time that the state had mandated a set of standards, the state legislature enacted laws requiring standardized testing for students at certain grade levels with the Oklahoma School Testing Program Act, Article V, Section 886, of Oklahoma School Law.

Oklahoma State Assessment Policies

The Oklahoma Secretary of Education, through the Office of Accountability had the task of seeing that school districts complied with the provisions of the Oklahoma testing act. The office of accountability would identify districts not making satisfactory progress and recommend corrective measures. Each year, results were to be reported to the public in individual school report cards that included test scores and other community information. In May of 1996, the Education Oversight Board was given complete control of the Educational Indicators Program. Information from the districts were gathered during one year, and since it was taken from more than one agency, some of the data might have been in conflict. The information would then be reported the next year, a full year after the fact.

Choosing Assessments That Measured PASS Standards

Oklahoma implemented standards developed by educators from the State Department of Education and classroom teachers. Conversely, the state assessments and testing program were controlled by the Oklahoma Oversight

Board, the legislature, and the governor's office. At the state level, standardized testing proponents included politicians, like Governor Frank Keating (2000) and other policy makers who believed students should learn the same sets of standards at the same grade levels and should be tested periodically at those levels to "measure" and "assure" learning (p. 3). With increased student learning the goal, these people called for high standards and a system of accountability that ensured educational reform. The incongruity was clear: practitioners developed the standards and politicians controlled the assessments.

In the 1990's, Oklahoma legislators implemented laws to administer a nationally standardized norm-referenced test at the third grade, and criterion-referenced tests at the fifth and eighth grades. The contracts for these tests were made with several different vendors, and almost every year teachers saw a new change in the testing program. The vendors included in1994-98, both Riverside and Harcourt Publishing; in 1998-99, Riverside and CTB-McGraw Hill Publishing; in 1999-2000, CTB-McGraw Hill; in 2000-2001, Riverside; in 2001-2003, Harcourt and CTB-McGraw Hill (see Table 1, page 15). Moreover, teachers questioned the validity as true indicators of student achievement in Oklahoma, when comparisons could not be made from year-to-year. In addition, teachers believed those tests did not accurately measure what the state said educators should be teaching according to PASS standards. Some educators (Popham, 2002), along with a few policy makers, believed that students did not learn at the

same rate and that learning was "developmental." These educators believed that generic "off-the-shelf" multiple-choice tests did not accurately measure what students had learned and that the test results unduly and inappropriately influenced education policy. Popham (2002) also believed that standardized tests should be used appropriately, to drive and improve classroom instruction. The questioned off-the-shelf tests were being used for high stakes decisions in Oklahoma when district leaders thought it was not appropriate to do so. Still, many believed that Oklahoma educators must be held accountable for the standards and the only way to achieve this was with a rigorous testing program.

In 2001-02, the Oklahoma legislature enacted an Academic Performance Index (API) that ranked districts by a specific formula for baseline data, and 80-90% of that formula was derived from the debated and highly questioned standardized tests. The API was formulated to meet the mandates of No Child Left Behind. However, the testing program had been constantly shifting until many believed the state achievement results could not be used to accurately measure schools or teachers. A debate surrounded the appropriate types of tests, the way they were used, and what other relevant information should be taken into account when making educational decisions.

No Child Left Behind Requirements for Oklahoma

When the No Child Left Behind Act (NCLB) was approved by the federal government, the basic process by which states were asked to improve

achievement and test scores for NCLB (2001) included (1) to implement learning standards in place in all core subjects; (2) to put assessments into place that would give accurate data on which to make decisions; (3) to make data-driven decisions from the assessment data; (4) to intervene where the data said instructional programs need improvement; and (5) to evaluate interventions frequently, still using assessment data, learning to strengthen instructional weaknesses.

In addition to the emphasis on test results for all subgroups of children, NCLB (2001) required all states to hire "highly qualified" teachers or those who were working toward that status. Under the legislation, "highly qualified" meant that every classroom teacher of core subjects by the end of 2005-06 must hold a bachelor's degree, be fully certified and licensed by the state, and had demonstrated knowledge of the subjects taught by taking a subject area test.

Oklahoma had in place a certification process for highly qualified teachers and was far ahead in the process with its PASS standards being in place and well accepted. However, the decisions surrounding Oklahoma state testing, such as which test company and which test to administer, was still determined by the legislature. Over the years, the testing program became so makeshift that districts could no longer use the testing information for comparative purposes from one year to the next because the state had implemented different types of tests almost every year, written by various testing companies.

A Washington-based group, Achieve, Inc., was commissioned by several Oklahoma state entities to study Oklahoma's standards and assessment programs. Gehring (2002) reported that Achieve found a "hodgepodge of Oklahoma tests that send a mixed signal" to educators. The study concluded that Oklahoma's program of assessment tests "could be undermining attempts to focus learning around the state's academic standards" (p. 18). The group further recommended developing a "coherent" assessment program "so essential information could be tracked year to year" (p. 18).

Upon examining the Achieve report (2002), the study, *Measuring up: A Standards and Assessment Benchmarking Report for Oklahoma*, was diagnostic in nature and focused on the state's education policies and practices. The study indicated that rigor of Oklahoma's core curriculum tests should be raised. A concern was also voiced about the state relying solely upon multiple-choice tests, limiting the kind of performance a test was able to measure, an objection raised by many critics over the last decade.

Problems using Oklahoma Assessment Data

Oklahoma schools could not measure progress from the third grade to the fifth grade because the test results were from two entirely different tests. By legislative action in Oklahoma, schools assessed third grade students with a norm-referenced test that many teachers did not believe highly correlated with the Oklahoma Priority Academic State Standards (PASS). Also, by legislative action,

schools assessed fifth and eighth grade students with a criterion-referenced test written specifically toward PASS standards, a test that was not nationally standardized and its many critics said was far too "easy."

The state-mandated standardized assessments implemented in 1994 in Oklahoma for the third grade were the Iowa Test of Basic Skills (ITBS) written by Riverside Publishing Company. According to Susan Newkham (2002) in the Evaluation and Testing services at the University of Oklahoma, the ITBS was created in the state of Indiana for the standards in the state of Iowa. That was quite an irony to teachers across the state of Oklahoma who used data from those tests for years, and instead of focusing on Oklahoma standards, they focused on another state's standards and/or test. Also, according to Newkham (2002) the research from the Evaluation and Testing Department indicated that Riverside Publishing's norm-referenced tests (ITBS) only correlated with PASS standards somewhere between 60-80 %. With the norm-referenced ITBS tests being used, state schools aligned their curriculum and taught to the items and categories used for that particular test instead of teaching to the PASS standards.

History of Oklahoma Test Vendors

Oklahoma first contracted with Harcourt-Brace in 1994, then to Riverside and CTB McGraw Hill in 1998, then to CTB McGraw Hill only in 2000, back to Riverside only in 2001, and finally to Harcourt-Brace and CTB McGraw Hill in 2002 to write the state CRTs, specifically written for PASS standards. However,

the state left in place the ITBS for the third-grade core curriculum except for one year, and deleted the eleventh grade geography (Table 1). For the year 2000, teachers again administered the norm-referenced ITBS in third grade and the criterion-referenced PASS written by McGraw Hill, but the legislature failed to fund tests for high school students, and the eleventh grade geography was dropped. The plan had been to implement end-of-instruction (EOI) exams for core curriculum at the high school level, but that did not happen right away because of lack of funding.

To complicate matters, the state awarded the PASS criterion-referenced assessments in the next year, 2001, to Riverside Publishing who had also written the ITBS, a different company from the year before. While McGraw Hill had implemented many application and problem-solving questions, Riverside still tested basic skills and knowledge.

To replace the eleventh grade tests, Oklahoma finally contracted for end-of-instruction tests in English II, United States History, Biology I, and Algebra I. State school districts implemented the English II and United States History exams for the first time in the spring of 2001 and field-tested the other two subjects in the 2001-02 school year. These EOI exams were to take the place of the eleventh grade ITBS, but the state legislature did not fully fund the exams to be given until the 2002-03 school year. The bids for new contracts for the 2002 and 2003 were awarded to two different test companies. The third grade assessment would be

the Stanford 9, a norm-referenced test written by Harcourt Educational Measurement; and the fifth and eighth grade tests would be criterion-referenced tests also written by Harcourt. The end-of-instruction tests for both years were awarded to CTB McGraw Hill.

Table 1: 1994-2004 History of Oklahoma State-Mandated Tests

	Third Grade	Fifth Grade	Eighth Grade	Eleventh Grade
1994-1998	ITBS (NRT) Riverside	PASS (CRT) Harcourt	PASS (CRT) Harcourt	PASS (CRT) Harcourt
1998-1999	ITBS (NRT) Riverside	PASS (CRT) CTB McGraw-Hill	PASS (CRT) CTB McGraw-Hill	PASS (CRT) CTB McGraw-Hill
1999-2000	No Test	PASS (CRT) CTB McGraw-Hill	PASS (CRT) CTB McGraw-Hill	PASS (CRT) CTB McGraw-Hill Geography Only
2000-2001	ITBS (NRT) Riverside	PASS (CRT) Riverside	PASS (CRT) Riverside	PASS EOI (CRT) Field Test - Riverside English II & US History
2001-2002	SAT9 (NRT) Harcourt	PASS (CRT) Harcourt	PASS (CRT) Harcourt	PASS EOI (CRT) CTB McGraw Hill English II & US History
2002-2004	SAT9 (NRT) Harcourt	PASS (CRT) Harcourt	PASS (CRT) Harcourt	PASS EOI (CRT) CTB McGraw Hill English II (writing added), US History, Biology I & Algebra I

When test contracts were awarded on August 30, 2001, the Oklahoma Superintendent of Education, Sandy Garrett (2001), announced on the OSDE website: "It is necessary that we hire a vendor we think can meet mandates in law and maintain high standards of accountability." The real problem lay with a testing program that placed so much emphasis on test scores considered by many educators to be inadequate and not highly correlated to state standards.

From information gathered in meetings with the School Improvement Advisory Committee at the Oklahoma State Department of Education (OSDE), State officials were pleased with the state testing program because it gave the state more than one source for accountability. However, OSDE officials were not pleased that local school districts could not use their achievement test data to compare instructional programs from one year to another. This was a direct result of nearly continuous changes in the types of tests and test companies from 1994 through 2004.

Test Results Decide Low-Performing Schools

Added to the problem of an unstable, if not inappropriate, testing environment, the standardized tests in Oklahoma were no longer just a way for local school districts to measure their student achievement. Under the guise of accountability, scores were used to determine low-performing schools in Oklahoma. When math and reading scores fell below what was considered the "norm" for a period of two years, the Oklahoma State Department of Education

could and would, by law, take over management of the school. Oklahoma School law stated:

> School sites which administer any norm-referenced test as required by the Oklahoma School Testing Program which had a student average score in the lowest quartile of Oklahoma students and whose student average score falls below the national average score shall be declared a low performing school. School sites identified as low performing for three (3) consecutive years shall be declared a high challenge school by the State Board of Education. Section 847C (Title 70 O.S. § 1210.541)

In addition to labeling a school "high challenge" with the threat of a state takeover, state test scores were used to rank districts across the state in the Oklahoma Academic Performance Index. Oklahoma's Governor Frank Keating (2000) in his Learnpower Program to Reform Education stated that schools must be held accountable for the teaching and learning process, and he proposed testing in every grade, using the tests to evaluate schools and programs and for a proposed merit pay system, a high stakes political debate in and of itself. The first step was the API that ranked districts across the state.

Overview of Oklahoma's Academic Performance Index

Beginning with the 2001-02 school year, the Oklahoma Academic Performance Index (API) from Oklahoma Senate Bill 810 went into effect,

ranking high and low performing schools and assessing what continued monies go with the ratings. Eighty to ninety percent of the API rankings were the scores from state standardized testing.

Encouraging tough, "rigorous" standards in the API to be adopted, Oklahoma Governor Keating (2000) stated that an enhanced testing program would focus on basic skills, and "coupled with more ways to evaluate schools," it would "give school patrons the tools they need to assure that learning is taking place." He believed that instead of filling out more forms, schools would be filling young minds with knowledge. In one sentence, the Governor proposed to give power, "the tools," to local school boards, but with standardized testing as more than half the measurement, those local boards would had little say in what tests were used or what and how they were measured.

Oklahoma state law required that students be tested at certain grade levels as an act of "accountability" to confirm that school districts were adhering to Oklahoma standards. The state had empowered the Oklahoma State Department of Education to let bids and contract for the best company to meet the needs of accountability. Problems implementing the policy within the OSDE and with the testing companies resulted in a true "hodgepodge" (Gerhing, 2002) of test data. The concerns that most people raised were how the state could use these test data to make important decisions, especially with dramatic changes almost every year.

Oklahoma state educators had tolerated the plethora of test information that could not be used in longevity studies. Teachers who administered an NRT could not see the progress that their students made in the fifth grade who took a CRT because the NRT results can not be compared to the CRT results. Feeling impotent to influence policy they focused on educating children within the constraints of a confused testing framework. However, with the right information, teachers could lobby for an appropriate, coherent testing program for the state. Table 1 showed two different types of tests and two different companies had contracts for state testing for the last two years. Educators could ask legislators as to how and why this was decided. If there were stability to the testing program, educators would then have had reliable data to evaluate programs from year to year.

Oklahoma Plans for Future Achievement Tests

The Oklahoma State Department of Education (OSDE) announced in the summer of 2003 plans for implementing reading and math tests from grades three through eight within the next three years. These tests would be criterion-referenced tests, and plans had been made to drop the Stanford 9 norm-referenced test at the end of the 2003-04 school year. However, until the CRTs were in place and data collected, that information could not be used by teachers until the 2006-07 school year. Geography alone was to be implemented in the seventh grade in

the spring of 2005; however, reading and math were not to be administered at that grade level until the spring of 2006, creating more confusion.

Significance of the Study

This project was significant because of the following problem statements:

(1) If the present state-mandated tests for NCLB were to be used to improve learning and classroom instruction, teachers should have had clear, consistent, reliable data. Most of Oklahoma's data could not be used in longevity studies to track student growth. In addition, results could not be used because of the constant changes in tests and test companies.

(2) If teachers were to make the underlying concepts of NCLB truly work and were to strive to improve classroom instruction, they needed training in using this hodgepodge of data. This training could only be achieved if by looking at what current practices were already working for classroom teachers.

(3) In Oklahoma, schools were tied to the Academic Performance Index and ranked according to test results. School officials would always want to be better than average, but with the current system of "ranking," someone would be at the bottom and someone would be at the top, with funding and ranking tied to unreliable data.

(4) Jobs and careers for administrators and classroom teachers could be in jeopardy because of being on a school improvement list, much like that of other states where test results were considered "high-stakes."

(5) Difficulties would multiply if possible "manipulation" occurred in testing situations because or the API ranking and funding. The manipulation could be in deciding which students could be left out of the testing, which students had actually completed instruction, or which students could be placed in special categories.

Research Question

Central to the problem statements above, the framework generated one overall research question: How had "highly effective teachers found methods to use Oklahoma achievement test results and other educational data to impact classroom instruction and improve learning that would ultimately result in "adequate yearly progress" for No Child Left Behind?

Scope and Participants of the Study

The scope of this study includes (1) a survey of 30 elementary classroom teachers, (2) in-depth interviews with six key elementary classroom teachers, (3) observations of the six "highly effective" teachers, and (4) artifacts or "documents" of data used by these six teachers from a district of 4500 students in Oklahoma. The six teachers interviewed represented those who were not only "highly qualified" under NCLB, but considered "highly effective," those who had prepared their students well for the state-mandated tests. Each year, the students had better than average results, no matter which test company wrote the test, what changes were made by the state, or what kind of students were involved. The

three teachers who taught third grade and administered the norm-referenced test had student results above the 70th percentile every year for more than five years. Two teachers who taught fifth grade and administered the criterion-referenced test had 90% or more of their students at the proficient level (70%) by test time, and the remaining fifth grade teacher had 80% of her below grade level students at the proficient level. These teachers represented the very best in the district and the ones who had practices already in place that could be replicated in meeting the mandates of NCLB. They were "highly effective" because they had overcome the effects of the students' prior knowledge, backgrounds, or economic status.

Statement of the Problem

Considering the variety of state and federal mandates over the past decade, many "highly effective" Oklahoma classroom teachers used state or local assessment data to inform instruction. With the state data questioned by many as inadequate and even inaccurate, teachers had found ways to use at least parts of the information they received. Their methods should be available for replication by others to meet the "Adequate Yearly Progress" mandated in the No Child Left Behind Act, 2001.

Limitations of the Study

The researcher, the primary investigator for this study, was a classroom teacher for 18 years and administrator for five years in the district studied. When seeking how other teachers use data to inform instruction, elements of subjectivity

might be present for one who had been well versed in the research of educational reform. The researcher was also the one who transcribed the information from the interviews and interpreted meaning from the patterns and practices that were shared by these teachers, but great care was taken to be open to all ideas, no matter how contrary to current research they might have appeared.

Definition of Terms

Ability Tests included such tests as the Otis-Lennon and the Naglierri to measure students' innate abilities to learn. They were used primarily for placement purposes, such as gifted and honors programs.

Achievement Tests referred to the many different types of tests that measure a student's acquired knowledge. They were usually multiple-choice and scored by a set of standards (standardized) and might be norm-referenced such as the Stanford 9 or criterion-referenced such as the Oklahoma PASS tests.

Alternate Assessments usually referred to those other than the multiple-choice standardized tests, such as portfolios and open-ended essay assignments.

Backloading referred to the process of finding out what the state achievement tests would measure and teaching to those items that were included. It had been generally agreed that if any standards were not tested, they might not be taught.

Criterion-Referenced Achievement Test results reported how well students had mastered a set of competencies or learned a set of skills according to set

standards and a pre-determined performance level. The scores were usually given as a scale score or performance level such as advanced, satisfactory, limited knowledge, and unsatisfactory.

Frontloading referred to the process of putting standards in place and then teaching to the standards. A test was then designed to measure standards, but it would have been impossible to include every Oklahoma standard on the test.

Highly Effective Teachers (for this study) had overcome the effects of the students' prior knowledge, backgrounds, or economic status. They were "highly qualified" under the NCLB mandates; in addition, year after year, they were the ones who had better than average test results from their students.

Highly Qualified Teachers, under the NCLB legislation, meant that every classroom teacher of core subjects by the end of 2005-06 must hold a bachelor's degree, be fully certified and licensed by the state, and had demonstrated knowledge of the subjects taught by taking a subject area test (NCLB Act, 2001).

Looping to the teachers in this particular study meant the looping, or constant reviewing of skills throughout the year. Once a student had learned a skill, oral reviews, chalkboard demonstrations, and some competitions were observed so that students would not forget a skill that was necessary for the next grade level.

Norm-Referenced Achievement Tests were designed to rank order students across a continuum of broad achievement with high to low achievers

(Stiggins, 2001A). Scores of the students were compared to a norm group. Students were then classified into ability groups that might need additional instruction, such as the gifted or remedial students.

A Performance Assessment would be a demonstration of knowledge through such things as portfolios, writing, research, or oral presentations.

Standardized Tests were created to be administered and scored under controlled conditions, meeting specified standards, by administrators, teachers, and monitor.

LITERATURE REVIEW

Chapter II

In the last decade, qualified educators as well as respected politicians had encouraged authentic assessment throughout the public schools of America, with popular items at educational conferences the various workshops demonstrating "how-to" implement authentic or "true" assessments. Real assessment to the experts included portfolios verifying students' works, essay exams requiring knowledge and analysis of content, and problem-solving activities designed to advance abstract thinking. All the while, with Goals 2000 and No Child Left Behind, more standardized tests, norm-referenced and criterion-referenced, were being implemented nationwide to measure and compare student achievement. Big money was being spent for those tests, $330 million in 2000, up from $165 million in 1996, according to Achieve Inc. (Spending big bucks, 2000).

The Arguments for and Against Standardized Testing

From one perspective, standardized testing gave educators, parents, and politicians a way of measuring learning of students; however, there appeared to be major concern about teachers teaching the test, classroom time spent on formal assessments and other negative effects. For educators, standardized testing was there to stay, especially with NCLB, and results would be used to determine which schools were "high-performing" and which were "high-challenge." According to federal guidelines with NCLB, how well students performed on tests

could determine which students would graduate, which teachers and administrators were re-employed, and which schools would remain open.

Advocates of Standardized Testing

Defenders of standardized testing believed comparing scores within a state or across the nation was the best way to provide accountability, to raise achievement levels, and to predict future success for students, especially for those from the economically disadvantaged. These advocates (Bush, 2001; Keating, 1999; Noyce, et al., 2000; Paige, 2002; and Riley, 2002) attached high stakes to these beliefs, using assessment data to evaluate teachers, programs, and schools. These supporters believed that to bring about true reform, decision-makers must rely on data, the numbers, and then an analysis of the correlations, and that the subjectivity must be taken out of the evaluation. The ones who defended the timed, multiple-choice process included those who truly believed in holding schools and educators accountable for doing what they said they were doing.

Studies suggested most officeholders at the state and national level advocated standardized assessments with Presidents at the forefront. Past President Bill Clinton promoted voluntary national tests having "positive consequences" for education (Mehrens, 1998). President George W. Bush with his No Child Left Behind (2001) promoted a national system much like the standardized TAAS testing in Texas. Proponents maintained the best way to

evaluate student achievement was to compare scores of all students allowing for a uniform set of standards, findings that could be used for high stakes decisions.

The current literature revealed the support for standardized testing came because of the following (not meant to be inclusive).

- Students who read poorly and could not perform simple math
- Complaints of students not prepared for the workplace
- College remediation rates, with students poorly prepared for college
- Implementation of basic and remedial classes in college
- Poor comparison of U.S. scores with other nations
- Low expectations for students, especially minority and poverty
- Accountability for the monies provided by the state and federal levels

Since the early 1960's, studies indicated that promoting accountability in school reform with annual standardized testing and publishing of test scores created very high stakes decisions, including (1) enormous amounts of monies in state and federal grants for intervention programs; (2) highly controversial debates about denial of diplomas and exclusion from activities; (3) removal of teachers and administrators from low-performing schools; and, (4) most serious of all, denied funding for schools who actually need more money to survive.

One analysis of a data-driven school (Noyce, et al., 2000) suggested that if we did not use the test data to figure out such things as gender issues in science and math, poverty issues in overall achievement, and the impact of professional

development on student learning, we would be working blindly when it came to reform. Noyce et al.(2000) reported that there was an "institutional willingness" to use numbers to reveal and predict patterns that could influence policy. Promoting a data-driven culture in a school, according to the Noyce et al. study, included identifying patterns of student performance, strong and weak, tracking data as to gender, race, economically disadvantaged, and progress for the limited English proficient (LEP). Educators would look at data to identify trends and suggest interventions to improve student performance. This information promoted using data to drive instruction, rather than being used to evaluate schools and personnel.

Most parents' concerns included their children not getting the best education they could possibly get. Also, parents wanted to protect children from failure, especially those from the disadvantaged, because they usually did not have a choice of moving or sending their children to a private school. Most teachers were concerned about students' learning and were doing the best jobs they knew how, but 69% of the teachers from a Phi Delta Kappa poll were opposed to Clinton's proposal for a voluntary national test (Mehrens, 1998, p 3).

A report from the National Assessment of Educational Progress (NAEP) showed that minorities, who comprised 40% of the nation's under-eighteen population by 2003, fared poorly in math, science, and reading. The College Board committed to spending $10 million to study and recreate educational

programs in the next few years for these minorities (Minority Students, 2000). But in many schools, even large districts such as Brazosport, Texas, with 14 thousand students, the tracking showed the correlations of low-performing students no longer linked to minorities, as the general assumption was once held, but to the economically disadvantaged. The poverty-level students were among the lowest performing students of all subgroups.

Defenders of standardized testing (Bush, 2001; Keating, 12000; Noyce, et al., 2000; Paige, 2002; and Riley, 2002)) also stated that poverty-level students were reasons for major school reform and those steps toward national standards of learning were imperative. Minority students were also of great concern to these defenders of improved education. The nation needed clear mandates to provide all students with opportunities to learn the same things higher achieving students in affluent districts learned. The argument posed by defenders was that with tough standards and high expectations, achievement would rise, and research showed that in some districts, those reforms produced higher results and better student achievement. When scores went up, administrators and politicians were happy and acted as if that meant achievement had gone up (Mehrens, 1998). If scores went down, disagreements would then arise over measures and cause-effects.

Critics of Standardized Testing

Critics of standardized testing (Berger, 2000, Bolon, 2000; Fair Test, 2003; Meisels, 2003; Popham, 2003; etal.) believed that test results were put

before the needs of students. Test results increased competitive pressures for students and parents, and teachers were under such strain to produce results that the process led to abuses of the system. These critics believed the consequences of large-scale assessment had yet to be proved and results were of little value. In addition, test development had been questioned for accuracy, relevance, validity and reliability. These reviewers believed very negative consequences could occur when real instruction was diverted with the decision-makers' fixation on test scores and high achievement. Parents' concerns were also voiced by these critics.

From this perspective, a parent (Zukas, 2000) reminded educators he did not wish "to sacrifice" his child on the "altar of the future." This parent preferred to keep politics out of decisions that would "create opposition" (p. 54). He would not get excited about the "latest" advances, claims of big performance gains, or schools forcing reform and his confidence in the compassion of the educational bureaucracy had been near zero. He wanted the educational community to "get it right" the first time because there would be no second chance for his child.

Highly respected researchers such as Robert Stake, University of Illinois, tended to agree with this parent. Stake (1998) reported that standardized testing had to do with "quality control" of schools. He said it had more to do with management, board oversight, parent complaints, school guidelines, and regional accreditation. Moreover, one of the complex and contradictory problems was defining what 'they" want of an educated person, and he had concerns over what

tests measured. He felt it imperative that educators not only examine for accuracy, relevance and freedom from bias, but that independent measurements would be used to confirm that scores indicated what we thought they should indicate.

Stake (1998) concluded that assessments did not indicate quality of teaching. He said there were real dangers and negative consequences for students, such as the following: real instruction being diverted, teachers intimidated, locus of control more centralized, undue stigma affixed to a school, falsification of scores, poor instruction, withholding of needed funds, and too much time away from instruction preparing to take the tests. On the other hand, he agreed that the state had a vested interest in what every child was learning and special attention needed to be given to the children who were least privileged and most vulnerable.

Testing Scandals

A news commentator reported that scandals in New York schools supported the testing critics' views and led to Chancellor Rudy Crew's contract not being renewed by his school board. From the *News Connection*, Branch (2000) reported that nine educators were allegedly involved in the scandal and at least eleven more faced similar charges. The educators had helped students cheat on tests by giving them answers or practicing test questions in advance. In some cases, wrong answers were erased and corrected. The reason given for the New York scandal was because teacher pay was directly linked to student performance. The concern was because politicians had promoted test scores being the primary

measurement of learning, but because politicians could understand and relate scores, they could make political arguments with the numbers.

Other testing scandals were reported over the last few years, such as one in Houston where a principal and three teachers were dismissed for prompting students during test sessions (Bolon, 2000). Most informed educators could cite instances where a state teacher or an administrator had had his/her certificate revoked because of teaching the test or changing answers.

Problems with Standardized Testing

There were problems with the process of standardized testing from the beginning (Berger, 2000). The following were some of the problems (not meant to be inclusive):

- Validity and reliability of tests and test scores
- Mandates from the top being a poor way to manage schools
- Real instruction being neglected
- Teachers and administrators jobs linked to performance (like coaches)
- Stigmas attached to low-performing schools and students
- Schools and teachers falsifying or manipulating test scores
- Time spent teaching test taking skills

One critic (Wallace, 2000) believed the achievement exams unfairly created undue pressure on teachers and students, and even worse, the tests failed to assess what students needed to know for the future, with some parents refusing

to allow their students to take the exams (p. 66). On the other hand, Mehrens (1998) believed the stress level for teachers increased, but he argued whether or not it actually increased the stress levels for students. He did not believe a major problem in education was students trying too hard to learn too much, a conclusion with which many classroom teachers would agree.

Educational Accountability

Because many people, politicians at the forefront, believed America's schools were doing a poor job of educating its children, schools strived to become better – more accountable. Leaders, President George Bush, Oklahoma Governor Frank Keating, and others, believed standardized "secure" testing was the way to accomplish the necessary accountability, so schools would continue to live with the consequences, both negative and positive.

Also, schools had a responsibility to see that students had been taught what was required for the test. School officials must make sure that students have had the opportunity to learn what would be tested, and classroom teachers must ensure that happens. In addition, educators should insist that standardized test data be used appropriately in state and federal policies. High stakes had been attached to test scores when a state ranked districts on an index and when they labeled struggling schools as "low-performing" or "high-challenge."

Test results could be used for accountability or they could be used to improve instruction. When respondents of a Gallup Phi Delta Kappa poll (2001)

were asked whether "tests should be used to determine how much students had learned or to determine the kind of instruction needed," the study stated that two-thirds of respondents indicated that tests should be used to drive instruction (p. 44). This finding may encourage those who believe that this was the basic purpose of testing. The poll offered less hope to those who believed tests should be used to support high-stakes decisions.

Could Test Policy Change Practice

Many teachers had viewed tests as a useless measurement of student learning, but the policies, standards and assessment, were designed to change that belief. If tests were designed for evaluation purposes, studies show it was very difficult for policy to change practice, especially when it involved state mandates. Proponents of using assessment data ((Bush, 200l; Keating, 1999; Noyce, et al., 2000; Paige, 2002, and Riley, 2002) believed that testing provided information about student progress that could be gained in no other way. They admitted that there were good sides to using the data, such as clearing up false assumptions and bad sides, including the lack of critical thinking tested. Concerns about teachers not going beyond what was on the test was a major concern for proponents, and the tests could affect what actually got taught.

Critics understood that achievement tests were there to stay, but they wanted good data used effectively, to improve instruction, to support professional development, and to change things in the right way. They concluded that if

assessment practices and policies were in place that provided reliable, valid data, the resistance to using the data might not be as great.

Current Rewards/Sanctions Provided by No Child Left Behind of 2001

In January 2001, the No Child Left Behind Act was approved by the federal government and required states to improve achievement and test scores. The mandates in NCLB were for states to

(1) develop and implement learning standards in place in all core subjects;

(2) develop and administer assessments that would give accurate data on the progress of all children;

(3) make data-driven decisions about school improvement from the assessment data;

(4) intervene where the data said instructional programs needed improvement; and

(5) evaluate interventions frequently, still using assessment data, striving to strengthen instructional weaknesses.

The Act also required states to prove "Adequate Yearly Progress" from one year to the next, with the goal of all students proficient in the school year 2013-14.

Adequate Yearly Progress

From the results of standardized tests administered annually, states were to develop target goals to make annual "adequate yearly progress" (AYP), becoming proficient in reading and math from the third through the eighth grades – all

according to the most recent reading and math test results. The scores were to be disaggregated into subgroups of poverty, race/ethnicity, disability, and limited English learners. Each subgroup would make AYP each every year until all students were on grade level and proficient in reading and math by the year 2014. Random sampling of groups of students with NAEP (2003) assessments in the fourth and eighth grades would provide a national measuring stick by which states were held accountable for rigorous content. NCLB required schools that failed to meet AYP to allow students to transfer to a better school, to provide supplemental services, to replace teachers and administrators who continue to do a poor job, and to restructure schools that stay on the "low-performing" list.

In addition, NCLB placed sanctions on states that failed to put into place standards and assessments within federal guidelines, allowing federal funds, 25%, to be withheld from those states who do not meet the deadlines. In contrast, NCLB could award states who significantly closed achievement gaps or exceeded AYP for two or more years in a row, but the Act did not make clear what the rewards would be or who would pay for them.

A school would be identified for school improvement if it did not meet AYP two years in a row. If the school continued on the improvement list and failed to make the improvements, the next steps were mandated by the state and federal governments. In year three, corrective action would bring about serious changes in a school or district.

Table 2: NCLB School Improvement Steps by Year

Year One	All students would be offered public school choice. Schools must receive technical assistance to identify and address problems, including analyzing data, improving professional development, using research-based strategies, promoting parental involvement and making a two-year school improvement plan.
Year Two	All students would be offered public school choice. Schools must receive technical assistance to implement their school improvement plan. Schools must offer supplemental educational services to students from low-income families.
Year Three Corrective Action Year	All students would be offered public school choice. Schools must receive technical assistance to implement their school improvement plan. Schools would be required to take corrective action to bring about immediate and meaningful change in their programs. The corrective actions must include at least one of the following: 1. Implement curriculum based on scientifically-based research 2. Replace staff deemed responsible for continued failure 3. Significantly reduce management authority 4. Extend the school day or school year 5. Appoint outside experts to help reorganize the school 6. Reorganize the school internally.
Year Four Restructuring Year	All students would be offered public school choice. Schools must offer supplemental services to students. Schools were required to prepare a plan to implement one of the following: 1. Reopen the school as a charter school 2. Replace the administration and staff 3. Reopen under private management or other governance 4. State Takeover
Year Five	All students would be offered public school choice. Schools must supplemental services to students. Schools must implement alternative governance of the school no later than the first day of school following year four.

By placing these sanctions on states, the federal government had made the multiple-choice reading and math achievement tests used for the Oklahoma Academic Performance Index (see Table 4) true "high-stakes" items and that put the tests' use and abuse in the middle of widespread dispute. Under NCLB guidelines, each state, including Oklahoma, was required to develop one single, statewide accountability system that measures AYP from state assessments.

Oklahoma assessments were to be developed according to state standards, and by 2005-06, the state had to administer the new annual assessments for grades third through eighth. Federal monies, $12.5 million for Oklahoma, were to develop those tests. States might defer only if Congress fails to appropriate monies. In addition, samples of fourth and eighth grade Oklahoma students would participate in NAEP testing to validate the state's reading/math tests. At least 95% of each student subgroup were required to participate in the assessment process.

New monies would be allocated for those schools not meeting AYP, but those funds would be dedicated for such things as supplemental educational services and professional development for "highly qualified" teachers. Schools meeting AYP would not receive additional monies. Publicity surrounding NCLB left the impression that schools would be rewarded for improving test scores, and the federal guidelines did allow states to reward schools who make their target goals. The irony that educators pointed out was that if the school was doing a

"poor" job of educating children, the school would get more money for resources under the Federal Title programs; and the school that was doing a "good" or "pretty good" job would not receive any new resources.

Highly Qualified Teachers

With the new "highly qualified" guidelines, teachers in all states must have held at least a bachelor's degree and have passed a "rigorous" state test in their subject area and educational preparation. To teach a Title I class which was underwritten with federal funds, all teachers must have been "highly qualified" by the end of the 2005-06 school year, but progress must have been started on those goals by school year 2002-03. Districts were allowed to use 5% of their funds to ensure the goal was met. Oklahoma had both requirements for certification since the early 1980's, but teachers who received their standard certification before 1983 and those who had been alternatively certified in certain years did not have the necessary qualifications. Federal monies were set aside in Title IIA for those teachers to gain compliance within three years, starting in 2002-03.

High Stakes Decisions for Oklahoma

The federal guidelines under NCLB had made student achievement test results a "high stakes" debate across the state. Oklahoma schools assessed third-grade students with a norm-referenced test that teachers knew did not highly correlate with the Oklahoma Priority Academic State Standards (PASS). Also, schools assessed fifth- and eighth-grade students with a criterion-referenced test

written specifically toward PASS standards, a test that was not nationally standardized, and its many critics pointed out was far too "easy." At the high school level, end-of-instruction (EOI) exams were developed for English II, United States history, Algebra I, and Biology I. By state law, any school whose student average score fell in the lowest quartile of all Oklahoma students would be deemed a "low-performing" or "high challenge" school. With NCLB asking states to implement a testing program for reading and math that encompassed the third through eighth grades, Oklahoma looked at a way to use the same tests that were already mandated and in place to suffice for the new system.

In addition to labeling a school "high challenge" with the threat of a state takeover, Oklahoma state test scores were used to rank districts across the state in the Oklahoma Academic Performance Index (see Table 5). Beginning with the 2001-02 school year, the API from Oklahoma Senate Bill 810 went into effect, ranking schools, assessing what continued monies would go with the ratings. Eighty to ninety percent of the API rankings came directly from state standardized testing, both norm- and criterion-referenced. In the year 2001-02, baseline data were established for the state from test results of the year before.

Oklahoma Test Contracts and Vendors

The state empowered the Oklahoma State Department of Education to let bids and contract for the best company to meet the needs of accountability. School Laws of Oklahoma (2002) required that students be tested at certain grade

levels as an act of accountability to confirm that school districts were adhering to Oklahoma standards. Problems implementing the policy within the OSDE and with the testing companies had resulted in the plethora of test data.

Contracts had been awarded to one publishing company and then to a different one the very next year, going from Harcourt-Brace to CTB McGraw Hill to Riverside and back to Harcourt-Brace. One year, Oklahoma did not have the money to test third graders, so a gap existed in the data. The third grade test, norm-referenced, was different from the fifth grade test that was criterion-referenced, making it impossible to track the same set of students' reading or math scores from third through fifth to see the progress being made, let alone what interventions might be needed. The concerns that most educators raised was how the state could use these test databanks to make important decisions, especially with the changes almost every year.

Moving Beyond Current Assessments and Improving Student Progress

With the reforms of the last two decades, achievement test results in some states had improved, but according to the National Assessment of Educational Progress, that improvement was not consistent with their findings. With the NAEP assessments, there was not the strong pressure to "teach the test" or "teach to the test." Oklahoma educators might have learned how to prepare students for the tests, rather than having made actual gains in learning, which left schools with

the question of how to move beyond the current pressure to prepare students to pass new state assessments and actually improve student learning.

Becoming Data-Driven

The current literature supported NCLB overwhelmingly in that educators should become data-driven and accountable in decision making (Noyce, et al., 2000). However, accountability loomed as a two-way relationship and those accountable needed the resources to accomplish the task – reliable stable, valid information or data to create the goals. If educators felt they did not have the tools to do the job, including reliable data and the necessary resources, they would logically ask how they could be expected to carry out the mandates from NCLB.

Since the state of Oklahoma did not have data from one consistent source that was longitudinal, covering a time period that gave good indicators of trends, educators needed more than what the state provided as "one-time" yearly achievement tests. First, training would need to be provided for classroom teachers so they would know which state data to use, how to use the data correctly, and which data were not consistent or reliable over time. Then, schools should locate teachers who were doing a really good job of using data and improving learning in their district and emulate their methods.

Assessments That Improved Learning

According to most curriculum experts, assessment that improved learning most often occurred during the teaching/learning process. Also, teachers who

assessed day-to-day and emphasized progress rather than failure increased students' motivation to learn (Stiggins, 2001B). However, assessment as a tool to promote student achievement was not often found in traditional classrooms, because teachers were prone to use checklists that indicated the material had been taught or they moved from one concept to another without letting the test results inform instruction. Rather than relying on worksheets or homework papers, teachers who used diagnostic and pre/post test methods of evaluation, including the student in every step of the process, would know exactly what students had learned and what they had not.

From diagnostic tests, educators could decide where students needed more practice and could revise their classroom instruction based on that information. At the same time the student learning was being evaluated, actual teaching practices could be evaluated, and a conscientious teacher would modify his/her instructional methods accordingly. If all students were not learning a basic arithmetical process such as fractions, the teacher could present the concept in a different way, specifically addressing all learning styles.

Assessments could be a simple process, such as demonstrations, and the assessments should be frequent. As teachers become open and forthright about their own work, that in turn, would let them know when to move on to the next concept. When teachers integrated and embedded assessment as a learning and

teaching tool, students demonstrated score gains on standardized tests (Black & William, 1998).

Involving students in the assessment process could increase the learning in a classroom as well as increasing scores on standardized tests. As students become involved, the tests would look more like teaching than testing (Davies, 2000). Students could be taught to self-assess, identifying the areas that needed improvement, setting goals based on the assessments and comparing their work to models of performance. Students could also be taught how to use rubrics with specific criteria to measure the quality of their work, and oftentimes they could be quite honest and even severe in their own appraisals. This method could provide non-judgmental feedback as opposed to student papers marked with red ink, and the teacher could learn to guide students in the self-evaluation process. Good methods of evaluation should be a natural part of teaching and learning, and good assessment also "takes time and is not cheap" (Sirotnik, 2002).

Purpose of Assessments

Most educators agreed that the main purpose of assessment should be to help students learn (Gallup, 2001). With positive feedback, students could find out what they learned successfully and what they had not. With tutoring, weaknesses could then be reduced. Many children could do well on assignments in class, but some did not, and if all students did not do well on a test, the teacher needed to stop and ask why. The assessments should be designed well within the

capabilities of most students, allowing for the regular experience of success, and children should only be assessed on knowledge and skills the teacher had given them opportunities to develop (Popham, 2002).

The assessment should always be based on curriculum objectives, but then again, no single method would give information on the full range of the teaching and learning of a classroom. If educators wanted a balanced, accurate picture of student learning, they should include alternative assessments (Bolon, 2000). These could be in the form of inquiry-based questions, open-ended questions/responses, structured-response answers, demonstrations, problem-solving activities, portfolios, projects and experiments. With the NCLB mandates, educators would have to prepare students for specific achievement tests, usually multiple-choice and still get the balanced picture of student learning within the same time frame each year.

Experts such as Bolon concluded that a responsible accountability system should be based on various, multiple assessments over extended periods of time, and many states do not have a stable, coherent system in place. Even Secretary of Education Rod Paige (2002) said, "Testing should not punish, but it should be an integral part of determining whether or not students were making progress." For Oklahoma educators to reach the point of truly using testing to make progress, the state would need to provide reliable, longitudinal data and allow districts to participate in self-assessment as well as state-assessment.

Use of Previous Oklahoma Test Data

With NCLB, the debate continued on how to use Oklahoma's standardized testing mandates, and many educators continued to view the tests as a useless but necessary exercise that interrupted valuable learning time. School districts could glean some test data to measure some learning objectives and still satisfy state policy makers, provided they could get the results in a timely manner. At the same time, schools could also continue to pursue assessment programs that would actually improve instruction, still questioning if the tests and test questions measured what they should measure and if the data gave teachers a focus on important basic skills.

Many Oklahoma school districts felt there were far too many state standards. Popham (2003) suggested that states should provide districts with a reasonable number of standards, along with assessments that measured those standards to be given at optional times during the year to determine the students' progress and programs' successes. With NCLB, Oklahoma school districts would continue to assess students with the same types of tests. With the overwhelming amount of time, effort and dollars that were being poured into the process, the test results should offer meaningful data that would actually be used.

Norm-Referenced versus the Criterion-Referenced Tests

The one factor missing from state-level discussions related to the different types of standardized testing. Because most of the tests were multiple-choice did

not mean they were the same. A norm-referenced test (NRT) compared how well one student performed on a multiple-choice, normed test with another population who took the same test. The NRT that Oklahoma chose to use did not correlate well with the PASS skills taught. State leaders and many teachers did not like the NRT because it simply compared how well one group of students did in relation to another group of students; therefore, the state began to phase in a criterion-referenced test (CRT) which was more diagnostic in nature. The CRT measured a percentage of the identified criteria, the skills and concepts that the student had mastered. Oklahoma's CRT's were standardized tests, but they were not nationally normed. Basically the NRT showed how a student did in relation to other students and the CRT showed how the student performed in relation to a set of criteria. Both were percentages, but many politicians and other people, teachers and educators included, did not understand the differences in those scores (see Table 3). For the 2000-01 school year in Oklahoma, if a set of students scored at the 70th percentile on the state NRT in math, they performed better than 69% of all other students. That same set of Oklahoma students who took the state CRT two years later in 2002-03 school year may have a score of 100, not indicating all made perfect scores, but that all passed the math portion at a satisfactory or better level (usually at the 70th percentile or above).

The type of scoring reported was misleading for many teachers, administrators, parents, and politicians who did not understand the educational

language. Even when the process was explained again and again, many still did not understand. Boards of education, which usually consisted of non-educational people, had a difficult time understanding the differences.

Table 3	Difference in NRT and CRT Scores
Criterion-Referenced Test	A 70% means that 70% of that group of students scored at a satisfactory or better level (the 70th percentile). A 100% means that everyone in that group scored at or above the satisfactory level (the 70th percentile).
Norm-Referenced Test	A 70% means that group of students scored better than 69% of all other students who took the test. A 100% would mean that group of students scored better than all other students across the nation (a much different score than the CRT's 100%)

To complicate the reporting, many people believed the NRT and CRT were diagnostic tools, and they did give valuable information, such as the rank order of students used to place students in gifted or remedial programs. However, the results did not measure how well a student could or did perform on classroom assignments or in work settings. With state administered standards-based tests, which might take two to three hours and that many sessions, educators did not get a true diagnostic picture. With a performance-based assessment, the administrator would give only enough of the test to determine the student's region of development, and a good administrator would stop the test at a child's frustration

level. After the information had been analyzed, the student could then be placed in appropriate, flexible groups for instruction until he/she was ready to move on.

The norm-referenced reports ranked how an individual or one group of students compared to another group. The scores were reported in percentiles as a measure of central tendency with 50th percentile as the median, or the point above which and below which half the national or "norm" group falls, thus producing a bell-shaped curve. A group of students could be ranked at the 70th percentile, which was twenty percentile points above the median, meaning that group performed better than 70% of all other students. That 70% may be an excellent score, depending on the criteria used and the difficulty of the questions. Students were measured against other students who fell within the same range of skills; they were not measured against a high standard of expectations. People who were not educators and not familiar with all different types of tests might assume, because it was the most common measurement for many years, that all standardized test scores were reported this way.

The criterion-referenced tests used the same multiple-choice format, but they were scored differently, comparing how well a student performed according to a set of absolute criteria, usually a preset blueprint. The NAEP assessment was a standardized criterion-referenced test where absolute standards were set and students' scores were then measured or compared to a perfect or exemplary performance. Usually, the scores would be in four categories, derived by stanines

and variously named, that would be akin to unsatisfactory, limited knowledge, satisfactory, and advanced, and it was quite possible for every student to make a satisfactory or better score on a CRT. There would not be a bell-shaped curve. Those scores were reported in percentages, measures of students who performed in a given proficiency rate or as satisfactory and above which was usually assumed to be at the 70th percentile. When this score was reported, educators and others should remember that the group who was reported at the 70th percentile was not twenty points above average. This score means that 70% of all students in that group made passing scores at the proficient rate (or 70%) on that test. The 70% score equates to a minimal or very average score.

On both norm- and criterion-referenced tests, a grade equivalent would usually be reported. Teachers and parents especially liked grade equivalents because they seemed to understand that concept better than a rank or percentile. A child in the fifth grade could be reported as reading at the eleventh grade level. This score was often misunderstood. What that grade equivalency reported was that an eleventh grader would have obtained that score on that fifth-grade reading material. It was unlikely that a fifth grader could manage eleventh grade reading assignments, but many teachers and most parents believed that to be the case.

Looking at commercial catalogs, approximately ninety to a hundred tests were available, including online assessments, and the quality varied, especially of those commercially prepared for large distribution. Because of demand,

publishing companies had begun to prepare open-ended and structured-response questions with the multiple-choice sections. They were more expensive than the multiple-choice and might not be practical for most school districts with the budget crises in almost every state. Critics had long promoted performance-based assessments, but portfolios and written assessments were difficult to assess accurately and cost-effectively. Several graders/evaluators might have to assess one students' work before a norm from a rubric could be reached, and even then, some subjectivity could still enter into the appraisal.

Key Components of the Oklahoma Academic Performance Index

In 2000-2001, Oklahoma Senate Bill 810 went into effect and the state legislature enacted an Academic Performance Index (API) as the accountability system that would measure the performance of all school districts in Oklahoma. The Index ranked schools from top to bottom according to a very specific formula and the Oklahoma State Department of Education (OSDE) would assess which districts were "high-performing" or "low-performing" and what continued rewards and sanctions would go with those rankings.

In the API, 80% to 90% of the formula would be from reading and math test scores derived from debated, highly questioned standardized achievement tests that Oklahoma administered at the third, fifth and eighth grades and to high school students with end-of-instruction (EOI) exams (see Table 4). For elementary and middle schools, 10% was added to the 90% reading and math

scores for attendance or school completion. For junior and high schools, another 10% was added to the 80% formula for an Academic Excellence Indicator. The Index was calculated mathematically to form a Baseline API from which targets were calculated. The amount of growth toward the target, or lack thereof, could be used for awards to schools that scored at or above the proficient level and sanctions for those that did not. The following lists a simplified version of items used in building the API Index:

Table 4: 2001 Academic Performance Index for Oklahoma

District Index =	80% reading and math scores from third grade norm-referenced and fifth, eighth, and EOI criterion-referenced tests 10% dropout, graduation, & attendance rates 10% ACT, AP, & college remediation rates
High School Index =	80% English II EOI criterion-referenced test scores (math not tested until 2003) 10% dropout, graduation, attendance rates 10% ACT, AP, and College Remediation rates
Junior High Index =	90% reading and math scores from the eighth grade criterion-referenced tests 10% attendance & dropout rates
Middle School Index =	90% reading and math scores from the eighth grade criterion-referenced tests 10% Attendance rate
Elementary Sites =	90% reading and math scores third grade norm-referenced & fifth grade criterion-referenced tests 10% Attendance rate

All 2001 baseline data were established from required district reports filed with the Oklahoma State Department the previous year. Attendance and dropout rates came from State Aid and Accreditation reports filed with the OSDE and were one year old, as were the graduation rates. The school test vendors provided the results collected from norm- and criterion-referenced tests administered in the spring semester the year before the baseline year. The American College Test (ACT) scores, Advanced Placement (AP) results and college remediation rates formed the 10% for an Academic Excellence Indicator used for high schools and were collected from the College Board, ACT, and the Oklahoma State Regents for Higher Education.

API Formulas and Growth Targets

State Statute 70 O.S. § 3-150 stipulated that certain indicators would be used in the API formula. Performance growth targets would be derived from the data, and a minimum growth target for each year would be used. The formulas would be outlined in a report that went to each school, but the information used to build the API formula was lengthy and very difficult for most educators and school administrators to understand.

While most classroom teachers and school administrators had no idea what the formulas meant, the Overview explained that "weights" were assigned and "multipliers' used to rescale the variables so appropriate ranges of values would be possible. An additional variable, the "reasonable limit value" or RLV

was selected based on historical performance patterns (see Table 5). For most people, those formulas would be quite intimidating, excluding the normal classroom teacher in making projections and hampering schools from using the information to improve or inform instruction. The following table lists a few of the formulas used in deriving the final Index number.

Table 5: Examples of Some API Formulas used in the 2001 API

Basic Block Formula : (CV – SM) * M + 1000	CV = Current performance value SM = State Performance Mean M = Multiplier determined by current state value, rescale outcomes, and the RLV 1000 = State Average Performance
Reading/Math Testing Index: (Q1 * 1) + (Q2 * 2) + (Q3 * 3) + (Q 4 * 4)	Quartile 1 = Unsatisfactory level Quartile 2 = Limited Knowledge Quartile 3 = Satisfactory Quartile 4 = Advanced Level
Deriving Multipliers: $\frac{\text{API Upper Limit (1500) – API Average (1000)}}{\text{Indicator RLV – State Indicator Value}}$	RLV = Reasonable Limit Value
Target API: CV API + (Maximum API – CV API) * .05) API growth points = (Target API – Baseline API)	CV = Current performance value 1000 = State Average Performance

In developing the formula, the API assigned for each school would be a single number and used to determine overall school performance and progress. Student information was to be disaggregated by student subgroups that included traditional and alternative education students. Also, the state, using federal guidelines, decided that data would not be reported for individual grade levels testing less than five students. Every element and indicator in the formula depended on accuracy in reporting from all sectors involved. Preliminary reports, with the information gathered, were sent to schools early in the year for them to check with all diligence. Many who received them checked to make sure the state had the right base information, but it was very difficult if not impossible to check the formulas. In addition to the formulas being difficult to understand, many believed the math and reading scores from the norm-referenced tests could not be used to accurately measure schools or teachers because the test did not highly correlate with Oklahoma Priority Academic Student Skills (PASS) standards.

Test Results Used for the Oklahoma API

When forming the Index, conversations arose among educators surrounding the appropriate types of tests, the way they were used, and what other relevant information should be taken into account. The testing program in Oklahoma had been in a state of constant alteration, of test companies and test instruments, until the data could not be used longitudinally. With the No Child Left Behind Act (NCLB) of 2001, states were charged with developing a single,

effective accountability system to measure progress of schools. The Academic Performance Index for Oklahoma, with a few minor changes, became the accountability system for NCLB. For the 2001-2002 school year, "new" Baseline Reports were furnished to Oklahoma Districts with "new" Baseline API data. Schools were told to disregard the 2000-2001 information because the data were not comparable in those reports.

Additionally, from results of assessments administered annually, states were to develop target goals to make annual "adequate yearly progress" (AYP) toward becoming proficient in reading and math from the third through the eighth grades. Oklahoma awaited federal approval before determining the AYP targets, which had originally been set at 5%. That target was not approved by the United States Department of Education and was revised into a formula that was difficult to understand. Even more changes had to be made when final approval was given by the USDE in May 2003. To meet AYP, test results were to be disaggregated into subgroups of poverty, race/ethnicity, disability, and limited English learners. Each subgroup would be expected to make AYP each and every year until all students were on grade level and proficient in reading and math by the year 2013-14. The random sampling of groups of students with National Assessment of Educational Progress (2003) assessments in the fourth and eighth grades would provide the national measuring stick by which states were held accountable for rigorous content.

Changes in Assessment Process

Most educators believed that to improve actual classroom instruction, the state of Oklahoma needed to completely overhaul its assessment process with new tests and a new system of accountability. February 19, 2003, the Oklahoma state superintendent from the Oklahoma State Department of Education (OSDE) announced that Oklahoma would receive a $1.4 million grant to do just that – to improve the quality of assessments statewide. Over the next three years, the state would lead a 16-state project to improve the system used to measure all students, including those with learning disabilities and those that were English language learners. Those tests that would measure student learning in grades three through eight would be developed within the next three to five years and then administered to measure state progress. However, those test results could not be used by classroom teachers until 2006-07, and it would be several years, three to five, before trends could be identified.

In the meantime, schools in Oklahoma were left with the same standardized achievement tests that they had administered for the last several years until they were phased out and replaced with new ones. To meet NCLB guidelines, Oklahoma used its Academic Performance Index already in place, with "minor adjustments" according to a letter from the OSDE, in its new NCLB application to the federal government. The OSDE felt it had developed a single, effective accountability system that measured AYP from state assessments.

In an open meeting at the State Department, a spokesperson for the School Improvement Division related that the state would have to change its five percent target goal because it would not meet the federal guidelines. A new application was filed by the OSDE February 1, 2003, outlining how the target growth from the API would be derived, but the approval was not given until May 2003. At the beginning of school in August 2003, schools in Oklahoma still were unsure how AYP would actually be figured by the state or how the formulas would be changed from the previous API. By March of 2004, a chart from the OSDE was received by districts giving specific targets for specific years in reading and math.

With NCLB asking states to implement the new testing program for reading and math that encompassed the third through eighth grades, Oklahoma assessments were to be developed that correlated to state academic standards. By 2005-06, the state would administer the new annual assessments for grades third through eighth and by 2007-08, it must add science with the reading and math. Federal monies had been allotted to develop those new tests, and the state might defer in developing the tests only if Congress failed to appropriate those monies. In addition, samples of fourth and eighth grade students would participate in NAEP testing to validate the state's reading/math tests. At least 95% of each subgroup must participate in the assessment process. Shaw (2000) believed that data warehousing could be invaluable for educators in making informed decisions about what and how they teach. When using data, Shaw reported examples of

teachers finding gaps in student success and the data showing what skills could be improved. With new tests, Oklahoma might be able to accomplish that goal.

In the meantime, Oklahoma administrators and teachers had to find a way to use the achievement data already in place until new tests were developed. Critics felt that the OSDE, even in the new system, would use the same tests and test companies that they had always used, plugging the numbers into the API system already in place while making only "minor adjustments." The implications were that teachers would be "teaching to the test" instead of the Oklahoma academic standards and that administrators would scurry, finding strategies and interventions to raise their API scores to keep off the school improvement list.

With the budget crises, teacher layoffs and other cuts, all over Oklahoma and the Nation, the gap widened in having the tools and resources for putting the correct interventions into place. Districts were cutting programs from their schools that contributed to a student's all-around education while making sure that reading and math test preparation gets funded. The federal government, and in turn state governments, had, in effect, seized control of education, raising the "stakes" to a very high level, bringing accountability programs to a crisis level.

National Context – How the Oklahoma API Compares to Other States

Some educators thought the guidelines for creating accountability systems in No Child Left Behind would bring about needed change. Others thought the

new laws would create major problems, even for "good" schools, including overcrowding and bringing test scores down. Many proponents thought the passing of rigorous state tests and mandates of AYP would benefit students, especially the minority and disadvantaged, while others saw the tests as unfair to disadvantaged students, causing major "dropout" crises. When President George W. Bush challenged the nation's public schools with NCLB, some cheered the mandates designed to bring equity for all students while other organizations reacted in dismay to the new regulations. In 2000, very few states had both standards and assessments in place that met the new regulations. Some states did not have assessments that matched their standards, and some had no assessments.

Monty Neill (2003) of the National Center for Fair and Open Testing reported that if the nation were to "leave no child behind," states must support higher-quality educational practices and use far more than standardized test scores to decide whether school districts had improved. Neill argued that continuing the course of using high stakes tests as the only element in the states' accountability systems would deepen the crisis in schools and cause even more dire consequences than those already endured.

FairTest (2003), Neill's organization, rated all states in five categories and only one received a five, as having a model system that performed well across all standards, including classroom-based assessments as part of the accountability system. That lone state was Vermont, and that rating system is outlined in Table

6. This group rated Oklahoma in Level two as a state assessment system needing many major improvements

Table 6: How FairTest Scored the States

Level 5	A model system. VERMONT
Level 4	State assessment needs modest improvement CO, CT, KY, ME, MO, NH
Level 3	State assessment system needs some significant improvements IL, KS, MD, MI, OR, PA, RI
Level 2	State assessment system needs many major improvements AR, CA, ID, IN, MA, MN, MT, NE, NV, ND, NJ,NY, OH, OK, SD, TX, WA, WI
Level 1	State assessment system needs a complete overhaul AL, AK, AZ, FL, GA, HI, LA, MS, NM, NC, SC, TN, UT, VA, WV
	Not Scorable DE, IA, WY

Vermont's Accountability System

In a press release, the Governor Howard Dean of Vermont, the model state in Table 6 above, reported his state's concern about the NCLB assessment process. Educators there felt they already had a good system of standards and assessment in place, including a model that included portfolio evaluations, and when the federal guidelines were issued, they considered turning back federal funding for Title I programs in the amount of $26 million (Keller, 2002).

Vermont Education Commissioner Ray McNulty said his state had been working very hard on school improvement, but that 38 of the state's 354 public schools were presently failing and they projected that 110 of their schools would be failing under the new NCLB guidelines (Education Assessment, 2002). However, Vermont faced a budget crunch, much like that of all other states, and could not afford to lose the federal money.

Since 1997, the Vermont state department had assisted schools in meeting the Framework of Standards and Learning Opportunities assessments, and priority schools had been given special assistance from state coordinators. Their accountability system included portfolios and other locally developed assessments, with much of the responsibility at the local level. Rep. Gaye Symington related that her state of Vermont had invested hundreds of thousands of dollars since 1997 to overhaul the accountability and finance laws, known as Act 60 (Sack, 2003). Vermont did not know whether the federal government would approve any parts of the efforts they had already made with its Comprehensive Assessment System (CAS). The Governor said the new reforms were "incredibly expensive" and would require starting all over again, adding that he did not think the cost to the taxpayers all across the nation had been considered (Keller, 2002).

Spokespersons for the governor's office and the state department agreed that Vermont could not do without the $26 million in federal monies; however,

the governor wanted those involved to measure the cost of laws from NCLB and the problems involved in the development of new assessments. These state officials of Vermont did not see the need of school reform in a different package with a different name (NCLB) and projected less equity for students under the new guidelines. If model states like Vermont projected more schools on the "failing" list and major problems in the transition phase, then other states could face the same or even more obstacles.

Nebraska's Accountability System

When states knew that standards and assessment would have to be embedded into a state curriculum, Nebraska wanted to combine both local assessments with the state tests, and state department directors reported the plan had been successful (Roschewski, 2003). Nebraska implemented a School-based Teacher-Led Assessment and Reporting System (STARS) as their answer to the NCLB Act.

Beginning in 2000, local Nebraska districts were required to measure standards in grades four, eight and eleven. Each district was allowed to select its own assessment tools, criterion-referenced, those most appropriate for its classrooms; the districts were then required to use the data to improve instruction. In the beginning, teachers spent many hours in professional development learning how to make appropriate assessments and how to use the results. The criteria established for the tests were (1) assessments should match the standards, (1)

students were given the opportunity to learn, (3) assessments were reviewed for bias, (4) tests were at the appropriate grade level (5) scoring was reliable, and (6) the mastery levels had to be appropriate. (Roschewski, 2003, p 518). From the first results reported, test scores were up for Nebraska, and if teachers truly used the assessments to improve instruction, then that state could have found the right combination to meet the mandates.

The Nebraska State Department of Education included in its guidelines that local assessments did not need to "add layers" to what teachers were already doing, but the tests needed to build upon assessment practices already in place. The assessments could include performance, observations, or paper/pencil methods (STARS, 2003). Nebraska's method would come more nearly meeting the federal requirements for testing all the standards. Too many times, states tested only a sampling of their wide array of standards in place. Instead of educators guessing at what needs to be tested, the Nebraska teachers were the ones who were directing what needs to be assessed and how well the students were progressing along the way.

Nebraska districts were then rated by the state with a Quality Criteria Rating Chart as to whether or not the district had met the criteria in a way that was exemplary through unacceptable. The raters who reviewed the portfolios included people from Nebraska or those who had had experience with that state's

educational background. If any of the criteria had not been met, the district would receive an actual feedback report instead of a number.

One of the items pointed out was that Nebraska's STARS did not produce a bell-shaped curve with some schools still at the bottom, but the tests were based on a criterion-referenced system that allowed all to rise to the proficient level. In comparing Nebraska's system with Oklahoma's API that relies solely on standardized achievement tests and ranks states above and below a mean score of 1000, much like a median on a norm-referenced test, it looked like a better system. If Oklahoma continued to use the API rankings, there would always be schools below the mean and always be those who say that standardized tests do not tell the whole picture.

As a result of the 1983 *A Nation At Risk*, every state except Iowa developed state standards and every state except Nebraska developed an assessment system. Because they had not had assessment policies in place before NCLB, educators there were able to look at many different systems to see what worked and what did not. What Nebraskans put into place looked very much like Vermont's system that combines local and state assessments, including portfolios, with much of the responsibility for assessment at the local level.

The Texas Accountability System

The assessment data from Texas, the "Texas education model," was reported in proficiency categories, and test scores had been widely reported from

many school districts as being much better since that state since had implemented the Texas Assessment of Academic Skills (TASS) assessments. Schools in that state published their test scores and proficiency status, especially those that were "exemplary," much like winning contests.

In a comprehensive study of Texas progress, Haney (2000) reported that the gains from TASS and the decreases in dropouts during the 1990s were more illusory than real and that the emphasis on TAAS was hurting more than helping teaching and learning in Texas schools, especially for at-risk students. From five different sources, Haney reported a sharp upturn in the numbers of students taking GED tests to avoid the state tests, with slightly less than 70% of Texas students graduating in the 1990s. Haney, from Boston College, reported that in one three-year period, the TAAS showed increases, but the TASP, a college readiness result, showed a sharp decrease.

Another study refuted the allegations in Haney's perspective about the Texas tests. Laurence Toenjes and A. G. Dworkin (2002) at the University of Houston stated that "Haney's Myth is a Myth." Using Haney's own methodology and statistics, Toenjes and Dworkin concluded that none of the improvements in the Texas exams had been shown a myth or fraudulently obtained and that Texas scores had increased and dropout rates had decreased. The authors alleged that Haney had a vested interest in proving that TAAS results were a "mirage," because he (Haney) had been an expert witness for the Mexican American Legal

Defense and Education Fund against the state of Texas, claiming that the Texas exit exam was unfair and discriminated against minority students.

After his first study, Walt Haney published several more "parts" to his original work, each part numbered, answering charges of Toenjes and Dworkin, as well as others. However, other researchers had also questioned Texas gains. One such study, on the Texas accountability system (Klein, etal., 2000), reported the TAAS test score gains were not confirmed by NAEP, that Texas fourth graders made smaller gains on NAEP math than on fourth grade TAAS math tests; that even smaller gains show up on the eighth grade NAEP in math; and that the racial gap in test scores between white students and students of color were widening according to NAEP scores. In response to Haney's studies, Camilli (2000) reported that the "miracle" in Texas looked much like the median elsewhere and that of 35 states and two districts, Texas was good enough to earn a rank of 17th. With many schools devoting a considerable amount of time toward TAAS preparation, these authors questioned the validity of the gains and cautioned against the danger of making decisions, rewards or sanctions of teachers and schools, based on test scores that could be inflated or misleading.

Amrein and Berliner (2002) analyzed the results of high-stakes testing from eighteen states, with Texas included. They reported no compelling evidence from the high-stakes assessments that gains in learning had taken place; moreover, they reported that testing programs from those states too often appear

"distorted and corrupted" (p. 56). The one objection they raised about the NAEP tests was not the test itself and what it measured, but that schools might be able to choose which students to include in their sample. For instance, districts might choose to exclude the limited English proficient (LEP) or special education student. Evidence from the NAEP study indicated that in eighteen states, student learning was indeterminate, remaining at the same level as before high-stakes policies were implemented, and in some cases, student learning went down.

William Bainbridge (2003), from the University of Dayton and SchoolMatch, reported that the most damaging fallout from the Texas tests was the false sense of learning. The reality was that in 1998-99, the Houston school system had 18,221 seventh graders and two years later, they had 9,138 ninth-graders in the same system. Test scores looked like they were on the upswing, but the attrition rate in the Houston system was 53%. The score gains did not include all the students who dropped out of school because of the tests.

Many teachers in Texas reported their schools could find ways to opt students out of taking the tests to make the scores look better, that test scores were artificially inflated through retention of students, and that the dropout rates had increased sharply in some districts. After having spent millions on a model assessment system in Texas, there seemed to be no clear indication that test scores, graduation rates, or learning had increased, just a general claim that test scores were up with no real evidence to support those claims.

Other States' Accountability Systems

Florida had implemented a statewide minimum competency exam that students were required to pass before graduating from high school (Amrein & Berliner, 2002). Early gains on the exit exam were used as an example of how an accountability system could improve education. As gains hit a plateau and dropout rates among minorities and the disadvantaged were discovered, Florida's accountability was discarded. Suggestions were that competency tests promoted low standards and widely perceived as "dumbing down." Governor Jeb Bush and other coalitions in Florida had asked the state to "soften" the requirements of the Florida Comprehensive Assessment Test to receive a high school diploma. Suggestions to consider were allowing students to pass the SAT, ACT, or other college entrance exams in lieu of the FCAT.

The Amrein/Berliner study (2002) also reported that in Arizona, students in poor and high-minority school districts scored lower than middle-class and wealthy students on that state's AIMS high school graduation test in 1999. From that state's reports, 97% of African Americans, Hispanics, and Native Americans failed the math section; a number of white students also failed the assessments. Ridden with problems, the state postponed that test.

Massachusetts began its accountability testing in 1998 (Bolon, 2002) with a "highly regarded, state-of-the-art program" (p. 2). However, from various news reports, schools throughout that state were facing public criticism for its

Massachusetts Comprehensive Assessment System (MCAS). Critics attacked the tests as "superficial, overly difficult, and a poor assessment tool" (Overview, 2003). Failure rates were so high that members of the Massachusetts Association of School Committees voted 137-30 to ask the state legislature to suspend the MCAS graduation requirement, but the state kept the tests. Even the president of Harcourt Educational Measurement stated that MCAS should not be used as a single graduation requirement; however, the Department of Education in Massachusetts continued to present the MCAS as a tool to get students to study harder and to improve test scores. All the while, many public entities in that state criticized the system as biased and unfair. The test problems had landed in the courts, with lawyers battling over whether the state could withhold thousands of seniors' high school diplomas because of the MCAS results. Assistant Attorney General Pierce Cray (Vaishnav, 2003) said the requirements were legal and that the state had pumped millions of dollars into the program to help students pass the tests, even after high school.

The New Hampshire School Administrators association estimated that the NCLB law would bring in about $17 million a year in revenue (Sack, 2003). In contrast, the state would incur at least $126.5 million each year in new costs. A spokesman for the administrators said their argument was not with the assessment system but with the federal government's shifting the costs to the states.

According to a study done by David Gullatt and Marlene Ritter (2002) for the American School Board, nineteen states ranked schools on their use of assessments as the major part of an accountability system. Three of those states used test scores alone to rank schools with another three states conducting site visits as a part of their accountability. Three more states factored in demographic information. They reported that no two states used the same process or same indicators to derive school scores. By the close of the year 2000, eighteen states had made graduation contingent on student test performance and an additional five states had made graduation contingent on the test scores (Olson, 2001). Only these nineteen states administered the tests that were required by the NCLB Act; therefore, the number of tests given combined from all other states had risen sharply and would continue to rise. In equal proportion, critics' objections may also increase.

In addition to the battles that raged in the states over the implementation of NCLB, harm from standardized tests, or more properly the test companies themselves, had been verified by many states. Harcourt Educational Measurement in 2002 made errors while grading the exams taken by Georgia students. The mistakes affected the placement by lottery of students into a prestigious school (Tofig, 2003). Students had been sent improperly to summer school in New York.

Other test scoring errors had been widely reported. When high stakes were attached to the results, such as diplomas and admissions to highly competitive schools of higher education, the lawsuits would continue. Test publishers battled more and more errors. At the same time, they battled to expand while satisfying the market and delivering more and more results on time.

METHODOLOGY AND DESIGN

Chapter III

Introduction

Just as there were many ways of learning for a student, there could be many procedures for undertaking research; above all, the choices of methods should match the information needs of the study or the research problem. In addition, educational research should begin with a solid framework and multiple data sources from which to collect information. Accordingly, a case study was chosen for this research that relied upon interviews, observations, document analysis, and a survey to address the primary research question:

> How had "highly effective" teachers found methods to use Oklahoma achievement test results and other educational data to impact classroom instruction and improve learning that would ultimately result in adequate yearly progress for No Child Left Behind?

Case study research was selected because the research question above focused on the way a particular group, that of highly successful and highly effective teachers, confronted specific problems. Practices from these highly effective teachers would provide insights into how educational progress for all students could be made, whether state assessment data or local diagnostic data was used to inform, or transform, instruction.

Case Studies

Case studies proved prevalent throughout research in the field of education. With a case study, the researcher developed an interest or research question that required insight, discovery, and interpretation, focusing on explaining a phenomenon rather than hypothesis testing (Merriam, 1998). The case could include a student, a group of students, a teacher or school, or a set of teachers or schools. According to Merriam, if there were limited numbers who could be interviewed, the phenomenon was then bounded and could qualify as a case. With a heuristic approach in a case study, the researcher would explain why an innovation worked and would then evaluate the innovation, thus increasing its potential applicability.

The case study did not command any particular or "set" methods for data collection or even for data analysis. According to Merriam (1998), "any and all methods of gathering data, from testing to interviewing, could be used in a case study." However, Merriam did contend that certain techniques were more likely to be used, such as the interview. Case studies presented knowledge that was concrete and contextual rather than abstract, offering experiences that would resonate with readers. The study should take the reader into a group's life, and the way the individual reports were compared and contrasted depended on the purpose of the evaluation. Once case studies had been organized and written, the content analysis would identify themes and patterns. The analyst would then

identify categories that emerged from cross-classification matrices, giving a more concrete quality to the information (Patton, 1987).

A case study seemed appropriate for a research question that would provide insight from one group's effective practices, efforts that deserved description and explanation. For this case study, a survey was developed and sent to all third and fifth grade teachers in this district, teachers who dealt with the third grade norm-referenced test (NRT) and the fifth grade criterion-referenced (CRT) tests, and 30 of 38 teachers responded. This survey offered information on how informed teachers were generally about No Child Left Behind (NCLB) and Adequate Yearly Progress (AYP).

Then, six highly effective teachers were selected to be interviewed from that group of 38 teachers who were surveyed. Those six were chosen from nine possible teachers whose students' exceptional success on the NRT or CRT could be documented. With one teacher declining to be interviewed because of time and personal issues, the top six were then selected for case study information. For the six teachers chosen for this study, their innovations contained merit that deserved to be highlighted, presented and explained so others could replicate their practices. These six were "connoisseurs" as Eisner (1998) called them, selected because they were "highly effective" teachers as well as being highly qualified under NCLB. As these teachers were interviewed, observations were compiled

into a set of field notes and teacher/student documents that would corroborate teacher practices were gathered.

Background for the Case Study

The background for putting a framework together for this educational study came from the federal mandates outlined in the No Child Left Behind Act of 2001 (NCLB). Oklahoma teachers struggled to meet the mandate of adequate yearly progress from NCLB with the Oklahoma state standards and assessment programs and policies used to meet AYP. Within this Oklahoma district, a few teachers proved to be quite successful in preparing their students for the state-mandated achievement tests, no matter what those tests were or which students were enrolled in their classes, hence the six teachers that were chosen for this case study.

The process by which the federal government had asked states to improve achievement and test scores for NCLB (2001) was (1) to put learning standards in place in all core subjects; (2) to put assessments into place that would give accurate data on which to make decisions; (3) to make data-driven decisions from the assessment data; (4) to put interventions into place where the data said instructional programs needed improvement; and (5) to evaluate interventions frequently, still using assessment data, learning to strengthen instructional weaknesses. This was the process that researchers behind NCLB had identified as

the way effective teachers and schools went from being low-performing schools to high-performing schools (Paige, 2002).

The information that emerged about the above process, from the highly effective teachers who were selected, included commonalities that other less effective teachers did not share. Effective teachers seemed to have had a set of rites and rituals that they undertook each and every year that made them highly effective or able to impact students' learning so successfully. The way these educators used the data, how and what they analyzed became the search for this study, to gain an in-depth understanding of how the classroom teacher used Oklahoma's assessment process and other data bases, intended for data-driven decisions and the meaning or learning for the students involved. This research was done through an initial survey, followed by interviews, observations, and collection of documents to determine what methods teachers used in their success.

Data Sources

Survey

A survey was first gathered from district teachers who had worked with the state-mandated tests most often, third and fifth grade teachers. According to Merriam (1998), "any and all methods of gathering data, from testing to interviewing, could be used in a case study." The teachers could answer the survey within a few minutes time, giving the researcher a great deal of valuable information gleaned within a short period, in this case, two weeks. According to

Sagor (2000), the survey instrument was probably the most popular method of collecting expeditious data from teachers. A survey (see Appendix A) was completed by 30 of the 38 third and fifth grade teachers in the district who responded. Those results were used to gain a general overview of teachers' knowledge of NCLB and Oklahoma's standards and assessment system. Information from that initial survey was then extended to in-depth interviews with six highly effective teachers, because research studies that used qualitative data could uncover the breadth and depth of actual practice within classroom settings. The responses to the statements made in the survey could be marked "Never, Almost Never, Some of the Time, Always and Almost Always," using a range of 1 through 5 correspondingly. The data from the 30 surveys were compiled and then charted with a bar graph for each question (see Appendix A), providing visual information on teacher beliefs about local, state, and federal mandates under NCLB.

This initial survey was the only quantitative data collected, and it provided valuable background information about how informed teachers within the district were on the new state and federal mandates. The survey results (Appendix A) were meant to provide the basis and direction for succeeding interviews. The first two questions of the survey asked general information concerning teacher knowledge of national goals and mandates in NCLB. The next six questions addressed state level decision making in Oklahoma that would lay a foundation

for teachers' use of state data, both standards and assessment information. Then, four questions asked about any direction or support received from the district about the use of state data in the planning of instruction. The last eight questions were directed toward the classroom teacher and his/her own methods and use of state standards and blueprints, pre tests and post tests, building and implementing diagnostic exams, and then planning interventions from the data gained. The answers from this survey would indicate whether teachers used state data, their own classroom data, or a combination of both (or neither) to meet the mandates of Adequate Yearly Progress.

Teachers who responded to the initial survey were assured anonymity, with the only identifier a question asking years of teaching experience. With 30 respondents, identifying which teachers had completed the surveys was not possible; however, four still chose not to give their years of experience. The questions asked on the survey were ultimately meant to authenticate the qualitative data gained from subsequent information in the six interviews.

Interviews

To understand the experiences of actual teachers' practices required exploration beyond the set of numbers from a survey, and the basic method or tools for this research were drawn from the actual experiences or interviews of six subjects. These interviews attempted to capture a picture of the process as it actually existed, and this study posed to capture the "quality" work that these

classroom teachers already had in practice. These were the roots of qualitative research that were derived "most directly from the ethnographic and field study traditions" and emphasized the "understanding" of certain behaviors within a context (Patton, 1987, p. 29). This study extended to observation of the real-life context, interviews with "how," "what," "when" and "why" questions (see Appendix P), and an analysis of the information collected.

Interviews began after project approval by the Institutional Review Board. Three "highly effective" teachers who prepared students for the third grade norm-referenced test (NRT) and three "highly effective" teachers who prepared students for the fifth grade criterion-referenced test (CRT) were chosen for the interviews, to gain and compare information for both types of tests. The interviews focused on how teachers used achievement data, both the state NRT and the CRT, and not an analysis of the actual test results. By legislative action in Oklahoma, schools assessed third grade students with a norm-referenced test, the Stanford 9, a test that many teachers did not believe highly correlated with Oklahoma Priority Academic State Standards (PASS). Also, by legislative action, schools assessed fifth grade students with a criterion-referenced test written specifically toward PASS standards, which was not nationally standardized and its many critics said was too easy. The research information was provided for these two grades that could not use the statistical information for comparisons of progress.

Teachers in the third and fifth grades had to work with two different state achievement tests, the Stanford 9 and the Oklahoma PASS, where the data were reported in two entirely different ways. Contending with this assortment of data, the six teachers who were selected for the study were ones who were not only highly qualified under NCLB but "highly effective" and true professionals in their field. Eisner (1998) described a "connoisseurship" that comes from participants who were more than just specialists in their field; they were "wise experts" who could discern the value of a product or process and provide reasons for their success or for any conclusions that were reached. They were the ones who worked every day with children and understood their craft "better than the PhD's" (Palmer, 1998, p. 7)

In-depth interviews with these six key third and fifth grade teachers in this district, teachers who were "connoisseurs" in education, were conducted like that of other case studies in research that provided important information and focused on specific educational issues. The interviews were audio taped and transcribed. The transcripts were then imported into NVivo, (computer-assisted qualitative data analysis software, CAQDAS) to find emerging patterns from various cross-references of patterns and themes. Participants were given their final transcripts of the interviews to check for accuracy, allowing them to clarify or expand upon topics covered in the interview. Participants were also provided with a transcript of the interview and asked to indicate anything that could not be quoted.

The criteria for the interview sessions were first coordinated with a set of standard questions (Appendix P) for which all case studies would be transcribed. As each case explained his/her method of gaining success, observations, along with artifacts or documents to represent the findings, were noted and then charted. All the key activities in the interview process, along with afterthoughts and even emotions were recorded and considered in the analysis and then synthesis of information. The participants in the research were teachers who successfully prepared their students every year for the state achievement tests, both NRT and CRT. These teachers were not chosen because of popularity or because someone "thought" they were good teachers, or were they from the "elite" schools. They were chosen because they had been successful and "highly effective" according to state test reports, in getting their students to demonstrate high achievement in reading and math. These successes were validated by group reports of students' achievement test scores, results that remained above average on a consistent basis year after year, with students above the 75th percentile on the NRTs and students at least 80% proficient on the CRTs. These results occurred no matter what abilities students possessed, what test was administered, or which test company or vendor had written the tests.

Participants in this study were asked to share their successes and their methods. They talked about their work and methods with enthusiasm. When the interviews were conducted, careful attention was taken to only record the

interviewees' information, as the recorder was concerned not to add to the data or intentionally leave something out. This attention to detail was what Merriam (1998) called the "ethics" of collecting the data and not filtering the information, and then removing any items of bias, or what she called "prejudices, viewpoints or assumptions regarding the phenomenon under investigation" (p. 158). Patton (1987) stated that closeness did not make a "loss of perspective inevitable" and that distance did not assure any measure of objectivity. The evaluator, in essence, had to get close to the subjects interviewed and build a rapport with trust so that, not only confidentiality would be maintained, but the interviewees would be more willing to share ideas.

These six participants of this study wished to remain anonymous and a signed, dated, consent form was collected from each of them. Narratives of each case study were carefully written using pseudonyms, followed by comments on links and commonalities. These transcripts were then imported into NVivo and coded into categories or themes. In summary, the following requirements were outlined for the participants in the interviews. They were teachers who had

- Taught third or fifth grade for more than three years;
- More than ten years' teaching experience in Oklahoma;
- Prepared students annually for Oklahoma state-mandated tests;

- Succeeded with student test preparation as verified by state-mandated achievement test reports provided by the district, students above the 75th percentile on the NRTs and students at least 80% proficient on the CRTs;
- Agreed to share their information about using Oklahoma state and classroom test data; and
- Shared a "passion" for student learning and "devotion" for their work.

All Interview data were transcribed, coded and analyzed with the use of NVivo (CAQDAS) and according to steps outlined by Merriam (1998):

1. Assemble the raw case data into categories and themes that captured recurring patterns;
2. Analyze the categories or themes that emerged from the case studies;
3. Synthesize information by writing organized theme narratives, providing a matrix of summary descriptions based on the qualitative data; and
4. Link the categories and concepts from each of the studies, providing the sources of qualitative data.

Observations

Observations were completed during and after the interviews, and the consideration was on what promoted students' learning and what data teachers used to gain that progress. From the observations, and the reflections that followed, emerged a set of field notes. That writing began in the form of brief

notes while in the classroom. Then full field notes emerged from immediate reflections, with both analysis and synthesis of the information. The field notes were kept in a file folder for each participant, in narrative forms much like that of a journal. In addition, the answers from the interviews posed even more questions, so informal visits were made to the teachers to gain additional information. The answers were captured in the form of short notes, including specifics about the documents and artifacts collected. In each case, the teachers were very receptive of answering informal questions because they had had the time to reflect and offer additional information. These notes were then merged with the field notes taken during the initial interviews. As much detail as possible was recorded, including descriptions of the data use and documents followed by observer comments.

Observational data about teachers' methods of using the state data were collected, transcribed, and then analyzed according to the guidelines developed by Merriam (1998) for case studies.

1. Verbal descriptions of the person and the activities.
2. Quotations and comments on the substance of what people said.
3. Observer's comments, identified separately in the narrative

The data from these studies would prove to be a valuable asset to probe and understand the particular processes on individual teacher differences or unique methods.

Documents

Teachers, during and after the interview process, provided supporting documents. Those documents included student artifacts, achievement data sources, including item analyses, teacher-made diagnostic tests, charts/graphs, and software assessments. Once the teachers produced the documents, their authenticity was assessed by asking the history of the document, their intention and their accomplishment, all of which was noted in the field notes/journals. Analysis of the documents and their uses followed and notations were made as to value, strengths and limitations of each one. Various documents were then selected to support teachers' themes that emerged from the interview data analysis, or documents that contained insights relevant to the research question. The researcher examined the teachers' words and attitudes as well as the accompanying teacher materials and student assignments. The data from the classroom artifacts proved to be appropriate data sources and revealed the nature of student achievement.

Data Analysis

The research data analysis ensued from the survey, interviews, observations, and documents, transcribed into a narrative form and then coded or labeled. This process, according to Boyatzis (1998), should be completed with "patience and determination," being extraordinarily careful to prevent premature theme identification. The management of this step was highly successful using

the computer-assisted software NVivo that grouped information into nodes or "meaningful units" as Merriam (1998, p 179) suggested.

Case studies should "reveal how all the parts fit together to form a whole" (Merriam, 1998, p. 6) and should strive for a depth of understanding from the participants' experiences. This type of study lent itself well to sorting through the experiences of these six "highly successful and effective" teachers to find what shared methods they had put into practice, methods that really worked in the classroom with students. That information became what Patton (1987) suggested was "information-rich" studies used to uncover and illuminate a process. Patton added that "purposeful" samplings would establish the commonalities from a "unique" or atypical group (p. 58), in this case a group of teachers who had experienced success in their endeavors of student achievement.

The interview transcripts were imported into NVivo so they could be coded into categories or themes. An analysis of the findings ensued, followed by an interpretation from the links or "nodes" as indicated by cross-referencing the material with the software. From twenty themes that originally emerged, the nodes were collapsed into seven major topics. These findings were then condensed into theme narratives followed by a documentational table or matrix to articulate the links found in the methods that teachers used and the sources of data, or triangulation, of the outcomes. The findings in the matrices were worded into clear statements that would give less successful teachers information to

emulate or follow in using Oklahoma data and in preparing students for Adequate Yearly Progress required in No Child Left Behind.

The next step was cross-referencing the data and developing theory, presenting it in a descriptive or narrative form (Patton, 1987), creating understanding from various links provided by the software. Making sense of that information was a challenging step, but the work was made easier with the software. With the information from transcribed interviews or case studies, commonalities were organized into categories to evaluate teachers' classroom activities and processes. With the software NVivo, clusters of data emerged from the information and connections began to develop from the themes or "nodes" that were coded. Inferences from the themes were then used to form the categories for theme narratives and the matrices of nodes or attribute that followed. The initial survey, interviews, observations could verify the research implications outlined in the matrices, and/or the documents and artifacts collected. As themes emerged from the data sources, they were labeled into nodes in NVivo and implications or conclusions were reached and verified with a matrix indicating sources of information, whether from the survey, key interviews, observations, and/or documents or classroom artifacts.

Transferability and Replicability

Regardless of the type of research, transferability and replicability were major concerns that could be addressed through careful collection, analysis, and

interpretation of the data. A survey, interviews, observations, and documents and artifacts were collected for this study, and each provided detail through which conclusions could be reached by the researcher. Merriam (1998) believed that this type of research must "ring true to readers, educators, and other researchers," and that other researchers should have "confidence in the conduct of the investigation and in the results of any particular study" (p. 199).

Triangulation

This study was conducted with careful triangulation of data to provide the validity and reliability that would be consistent and gain the confidence of others. One particular study by Anfara, Brown, and Mangione (2002) entitled "Qualitative Analysis on Stage: Making the Research More Public" offered strategies they used in working with qualitative researchers. In order for those who do qualitative research, Anfara, et al., outlined methods that provided adequate and clear justifications for findings and conclusions from qualitative data. Although triangulation and other qualitative strategies were mentioned in research articles, rarely was any evidence provided to show how that was achieved. The Anfara, et al., article commented on the debate surrounding "rigor" and analytical defensibility of the qualitative study, and their strategy was documentational tables, or illustrations that assessed the research quality and rigor. A sample of the documentational table or matrix used at the end of each theme narrative is as follows.

Sample Table: Matrix of Findings for The Issue of _____

Theme 5: (Findings and Sources of Data) Sources of Data: I = Interviews, S = Surveys, O = Observation, D = Documents	I	S	O	D
1. Clear Statements derived from the qualitative research were made in the first column and then coded in the columns to the right with an "X" indicating the sources of information.				
2. The notations at the end of the table include the document or artifact used for the research and where it might be found.				

Notes: Document/Artifact for #1 can be found _____.

Document/Artifact for #2 can be found _____.

For this study that attempted to uncover the practices of "highly effective" teachers in using state standards and assessment data, the documentational tables, or matrices, were used, much like those suggested by Anfara, et al. (2002). Research conclusions were recorded after the survey, the interviews, and the observations had concluded. As themes emerged from the NVivo software, the exact participants were notated for each theme or statement. For example, a research answer or statement was numbered and tabulated followed to the right by the specific participants (i.e., P2, P3, P4, P5) who had agreed to the statement.

Of most value were the matrices that recorded the statements and research data in a table with the source of the data listed to the right, with an "S" coded for

survey data, an "I" for interview data, an "O" for observation data, and a "D" for document/artifact data. This provided four sources of data collection to corroborate the information gathered from the various sources. As Anfara, et al., (2002) suggested, this enabled the researcher as well as others to regard the material critically, "to test it, to identify its weaknesses, to identify where to test further doing something different" (p. 33). As the themes emerged, if they could not be corroborated with at least two sources of data, they were not included. On the other hand, the use of documents and artifacts to support the evidence added a third source to provide triangulation of important themes and statements that might not have been included otherwise.

The matrices clearly outlined the sources of data in documentational tables and added strength to the teachers' arguments as well as providing the transferability and replicability needed for this qualitative study. Each theme was presented in narrative form, followed by a detailed table of how the research findings were related to the data sources and how the triangulation was accomplished. This method answered the challenge of "trustworthiness of qualitative research" that Merriam (1998) described in providing the validity of interview information and observations in qualitative research (p. 203). The internal validity or credibility and the external validity or transferability, along with the reliability or dependability of the study, was accomplished through the documentational tables.

The Researcher

As an educator well versed in the education process for 24 years, I felt I was well prepared for this study. I taught language arts for 18 years, grades 7-12, including Advanced Placement, regular, and remedial classes. Elective classes taught included creative writing, journalism, yearbook, speech, and drama. Serving as a high school assistant principal of 900 students for one year, my duties included attendance, discipline, and formal observation/evaluation of teachers. For the past five years, I served as K-12 district Director of Curriculum and Instruction, in charge of many programs, such as testing and assessment, gifted and talented, all Title programs, professional development, and federal and state grants. With classroom experience at the secondary level, I had made a very concentrated effort over the last five years to understand what really happened in the elementary classroom; in the process, I came to admire the organizational and creative skills of the elementary teacher. Elementary teachers were very child-oriented, while the secondary teacher tended to be subject-oriented. When the time came to find answers for this district's teachers for NCLB, I looked to the elementary teachers; primarily because several of them had already found the solutions to meeting the NCLB mandates that others could emulate. My goal was to capture the practices of these highly effective teachers for others to follow.

In a concentrated effort, I was open and receptive, as outlined by Moustakas (1990) to the essential ideas being uncovered, providing illumination

from the participants' knowledge. Once the explication process began, that same openness and reception was used to discover hidden themes from the phenomenon. Creative synthesis, according to Moustakas, could then be achieved through those same intuitive powers. Any bias about what a good teacher should be doing with his/her data and all ideas of what methods teachers could have implemented were set aside to allow the experiences to offer the ideas. Merriam (1998) outlined steps for ensuring validity and reliability of the information that was gathered with multiple sources of qualitative data that included interviews, observations, artifacts, and a general survey. For internal validity, I compared and contrasted the findings to ensure they matched reality, and for external validity, the data were analyzed to see how it could be applied to other situations. With this study, I worked diligently to maintain both internal and external validity.

Community Context of the Research

Elementary schools located in one southwest Oklahoma public school district were selected as the sites for this research. This community of 24,000 people with a military installation brought a diversity of cultures into the school system. The elementary schools were 58% White, 24.4% Hispanic, 13.4% African American, 2.2% Native American, and 2% Asian. The district had a total student population of 4500, with approximately half as elementary students. Over 59% of the elementary population and 47% of the total student population lived in poverty as defined by the requirements of the federal free and reduced lunch

program. Of the total enrollment, 18.6% were eligible for gifted/talented programs, 10% eligible for special education services, 11.2% were bilingual students, and 1.8% had a limited knowledge of English. The district employed around 360 teachers, counselors, and librarians. With impact aid from the military installation, resources and salary were not large issues for the teachers interviewed; however, one expressed an expectation to have all of her health insurance paid. This district had placed an important focus on student achievement over the last five years, putting into place total curriculum alignment and a pre/post test process. The district had been awarded a Quality Award at the commitment level by the Oklahoma Quality Award Foundation, Inc., the first public school in Oklahoma ever to be honored with this state recognition.

Communicating the Case Study

The current literature overwhelmingly supported the need for change in education, supporting the mandates of NCLB (2001), in that educators should use assessments to improve learning and they should become data-driven in their decision making. The know-how, including the resources, to reach that stage had been missing for some Oklahoma educators who had the challenge of carrying out the mandates of NCLB without good data from the state's standardized achievement tests. What was not known was what Merriam (1998) called the "gap" in knowledge base and the problem statement. Therefore, the above "gap" was the framework that generated this research question and the methodology.

In the search for answers, teachers who had been concerned about student learning also commonly believed that Oklahoma's state-mandated testing system was a mixture of unusable information, often giving inadequate and inaccurate results. The state had contracted with one vendor and then another (see Table 1) and changed the exams from norm-referenced to criterion-referenced until the data could not be used in longitudinal studies. In case studies, analyzing highly effective teachers' methods of using data was one way to give all educators new insights into the problem of using Oklahoma achievement results for making the required AYP for NCLB.

The case study proved to fit the specific research question and subsequent answers that this researcher uncovered. According to Majchrzak (1984), the last step of the research was to "communicate" the ultimate results in a narrative form. That step would be completed outlining recommendations and implications of the case study through theme narratives. In qualitative research, the findings of the case study would be used to make improvements in the practice of others (Bogdon & Biklin, 1998), and that was the goal of this research. The interviews, observations, documents, and surveys with these classroom teachers who continued to teach with great passion and devotion or a "Heart of Hope" (Palmer, 1998) were those who had a message to share and could truly help make improvements in the practices of others.

RESULTS OF RESEARCH

Chapter IV

Introduction

This case study research was conducted in an attempt to capture proven practices of "highly effective" elementary teachers in one southwest Oklahoma school district. These teachers were chosen because they were not only "highly qualified" under the No Child Behind Act of 2001, but they were highly successful and effective in having their students ready for the Oklahoma state-mandated tests each and every year. This research section will report (1) information obtained from an initial survey of 30 third and fifth grade teachers; (2) qualitative data from interviews of six key or "highly effective" teachers; (3) observations of six key elementary teachers; and (4) documents or artifacts collected from these six teachers' classrooms.

Information Obtained from Initial Survey

An initial survey (Appendix A) was dispensed to all third and fifth grade teachers (38) in this Oklahoma district in order to obtain a general overview of elementary teachers' knowledge of NCLB and how to meet the state and federal mandates. From this survey, 30 responses were gathered. The respondents' years of experience ranged from 3 through 36 years for 26 of those teachers (4 surveys did not respond to this question). The mean years of experience for these teachers averaged 16.5 years and the mode equaled 17 years.

Summary of the Surveys

The first two questions of the survey (see Appendix A) asked teachers about their involvement in the mandates of NCLB. The 30 surveys indicated that this district's teachers believed the national leaders had left the Oklahoma classroom teachers out of the planning process when the state and national goals and mandates for NCLB were implemented (see Appendix A1). They also indicated that the federal government had not provided adequate resources for the implementation of the goals (see Appendix A1). Only a few teachers answered "some of the time" to both of those questions, indicating that mandates from above and the federal government in particular did not involve teachers nor provide the necessary resources, namely the funds, to meet AYP for NCLB.

The next set of questions on the survey, three through eight, asked teachers to respond to issues involving Oklahoma standards and assessments (see Appendix A). More teachers believed that Oklahoma involved its teachers in setting the Priority Academic Student Skills (PASS) that those involved in the federal mandates (see Appendix A2). Half of the respondents believed that Oklahoma had provided adequate, appropriate resources to teachers for NCLB (see Appendix A2). Almost half indicated that Oklahoma had well-defined clear standards in place for implementing NCLB (see Appendix A3). Only half or less of the respondents believed that Oklahoma had adequate assessments in place for NCLB (see Appendix A3); that the state provided the test results in a timely

manner (see Appendix A4); and that test reports were easy to read and understand (see Appendix A4).

The survey questions about the local district, nine through twelve, indicated the teachers felt they had a well-defined curriculum, had adequate district assessments, and had looked regularly at state test blueprints and achievement results (see Appendix A). The majority (23) of the teachers felt the district had a well-defined curriculum that matched the state PASS standards (see Appendix A5). The majority (25) indicated their local district had adequate appropriate assessments that provided data to meet the AYP for NCLB (see Appendix A5). Their district always or almost always (25) examined the results of student performance from standardized state tests (see Appendix A6). The majority (25) indicated that the district regularly examined results of student performance from district teacher-made tests (see Appendix A6). An interesting contrast existed as teachers indicated they planned instruction from district achievement test data more than they did the state data.

The next set of questions, thirteen through twenty, asked the participants to respond to local or campus level activities (see Appendix A). As classroom teachers, 22 of them kept abreast of the content standards and the current state blueprints that were provided by the Oklahoma State Department of Education (see Appendix A7). The one item that revealed a great diversity among the teachers was in using a pre test before they started a new unit of classroom study.

Only half of the respondents indicated that they pre tested their students before beginning a new unit of study such as fractions or punctuation (see Appendix A7). The majority of respondents (25) knew how to build diagnostic assessments that would give them instructional information about their students (see Appendix A8). The majority (27) also planned instruction according to the Oklahoma PASS standards just as the researchers behind NCLB showed that helped schools move from low-performing to high-performing (see Appendix A8).

When teachers were asked if they planned instruction according to the Oklahoma achievement test data rather than the standards, a wide range of answers were given (see Appendix A9). Eight indicated they always planned using the Oklahoma data, ten said almost always, ten answered most of the time, and two almost never, indicating that they used the test data only some of the time. That same diversity was found when teachers were asked if they planned interventions according to the Oklahoma achievement data; four answered always, fourteen said almost always, ten said some of the time, one answered almost never, and one said never (see Appendix A9). In contrast, teachers planned interventions according to district achievement tests more than they used state achievement data. All of teachers used the district data some of the time, with eleven almost always using the district data and eight always using the data (see Appendix A10). The teachers indicated they used the district and their own

classroom data more than they used the state data. From the information gained on the survey, the questions for the interviews were formed (Appendix P).

Interviews with Six Key Elementary Teachers

This case study research was undertaken to capture the proven practices of highly effective teachers, particularly that of teaching with data in mind, both state and local results, to plan instruction and to plan for the state-mandated achievement tests. Information from these interviews were transcribed and imported into NVivo software where the data were grouped into "nodes" or themes. As these themes or categories emerged from the interviews, they were cross-referenced into groups that provided the narrative by which the data would be reported. A total of seven themes were identified that directly related to the research question. They were (1) Oklahoma State Standards; (2) Oklahoma Assessments/Blueprints; (3) Evaluation and Use of Data; (4) Classroom Interventions/Benchmarks; (5) Time for Planning; (6) Discipline and Parental Involvement; and (7) High Expectations for Students. Also, further topics, or other issues or non-issues that effect student achievement, emerged from the interviews, relative issues minimally addressed by the respondents.

The above themes were supported from the initial survey of 30 teachers; (2) the transcribed interviews of six key elementary teachers, (3) the researcher's field notes and observations, and (4) supporting teacher and student documents

and artifacts. Following each theme discussion will be a summary statement and a matrix that indicates the source of supporting data.

Introduction to Six Key Teachers Interviewed

All of the participants in this study wished to remain anonymous and pseudonyms were used. Information was reported to erase any identifiers in the descriptions that followed. This list started in order of years taught in Oklahoma.

Table 7: Years and Grade Levels of Teachers Interviewed

Years Taught In Oklahoma	Years Taught at Present Grade Level	Grade Level of Teacher	State Test Given At This Level	Documented Results from NRTs or CRTs
11	4	Fifth	CRT	80% of students proficient
21	11	Fifth	CRT	90% of students proficient
21	20	Third	NRT	Students above 75th percentile
24	3	Fifth	CRT	90% of students proficient
31	26	Third	NRT	Students above 75th percentile
36	35	Third	NRT	Students above 75th percentile

Participant #1 (P1), Mrs. Patricia Arbello

Well-known throughout the district as an excellent teacher, Mrs. Arbello (Interview, 10/03/03) mixed many of the old teaching methods such as "drill/kill" with newer methods such as "hands-on" activities that were more enjoyable and seemed to motivate students. Mrs. Arbello was one of those remarkable teachers that liked all of her students, not just some of them. This was a trait that endeared

her students to her, from the very beginning of the year. If the students were a little gloomy on a particular day, she acted inane with them. On the other hand, she also discouraged them when they became too noisy or loud, with just one hand held in the air. The students knew that one hand in the air meant to be quiet and stay focused. She liked to make up amusing characters and stories to personalize her lessons and make them more interesting. She seemed to love her children more than most teachers, and she shared that as her secret weapon. She thought any child would learn better if he believed his teacher felt he was special, so she treated all of them as if they were the most special children in the world (Field Notes, 10/03/03). Well-organized and prepared, Mrs. Arbello knew exactly what she should be teaching. She used her PASS standards, curriculum guides, and test blueprints; moreover, she was the only one of the six participants who used state achievement data in depth. However, she did not think the NRT that Oklahoma had chosen for the past three years (SAT9) matched the PASS standards very well. She found teaching to the standards and then teaching to the test frustrating at times because she felt herself working in two different directions (Field Notes, 10/03/03). Every year, Mrs. Arbello made a list of skills that were tested from the NRT results she received at the end of the year. She used her summers to chart those skills and put them in order of "weak to strong." She began her year with the weak skills and concentrated on them all year, continuously reviewing and "looping." Then she introduced each of the skills on

the list one at a time, reviewing those and going more in depth. She liked to keep data, using a timeline to keep herself on track. Every year, she would worry about the test scores for her children, and every year, they scored extremely well, always above the 75th percentile on the NRT (Field Notes, 10/03/03).

Participant #2 (P2), Mrs. Sheryl Blair

A third grade teacher, Mrs. Sheryl Blair (Interview, 10/06/03) prepared her students well every year for the state achievement tests, and she disciplined her students in very appropriate ways from the very beginning of school. Mrs. Blair's students had scored above the 75th percentile on the third grade NRT every year for the past five years. She felt this discipline was the key to this student learning. Mrs. Blair said, "In the beginning, teaching was not so stressful. The biggest thing that has changed in all these years is the discipline. Children are a lot different than they were 30 or 40 years ago" (Field Notes, 10/06/03). This teacher's best quality was her devotion to her work; however, she appeared to have some frustration in meeting NCLB, the state mandates, and preparing students for the state test. She found her work difficult with the discipline that she had to establish every year, with several special education students mainstreamed into her classroom. She said those students took up an enormous amount of time, and she said they took time away the time she needed with all the other students to get them as far as she could. Although she was frustrated preparing students for the nationally normed NRT test, she did an excellent job of teaching and

disciplining the children, using a profusion of hands-on materials and the students demonstrated their learning in very creative ways (Field Notes, 10/06/03).

Participant #3 (P3), Mr. Steve Callahan

This participant (Interview, 10/15/03) was a no-nonsense instructor who worked very hard, all day, every day, and his students did the same. Mr. Callahan had very high expectations of his students; however, he seemed disturbed about the lack of prior knowledge that those students came to him with every year. This teacher felt that the lower grade teachers could have prepared students better in reading and in math (Field Notes, 10/15/03). There were parents who requested their children be in his classes because the students stayed on focus, learning and growing every day. From observations, he taught the basic state standards in all core subjects with ample guided practice (Field Notes, 10/15/03). Mr. Callahan prepared his students well for the state test, even though he felt the NRT did not highly correlate with the PASS standards. His students' scores were always at or near the highest scores, compared to all other third grade groups in the entire district. The one personal attribute he had that was different from other teachers was a commanding voice, an advantage in discipline for him. The students paid attention, and they respected him. His commanding manner was much different from the other participants, but it worked very well for him. His students scored above the 75th percentile year after year, with many children at the advanced level (Field Notes, 10/15/03).

Participant #4 (P4), Mr. Paul DiCarlo

Mr. DiCarlo (Interview, 10/27/03) was the most attentive and perceptive teacher this researcher observed. He knew his students extremely well, having pre-tested them at the beginning of the year and with each unit, using data to drive his instruction. Mr. Dicarlo's perception guided the students' practice, and as he watched them do their work, he asked them why they had completed each step of their process. If a student had made an error, he asked the child to come and sit beside him at his desk. He then led the child through the process again and again until he had corrected the error in the process (Field Notes, 10/27/03). Mr. DiCarlo said that it was difficult for a child to "unlearn" a method that he had learned incorrectly, and if teachers could be observant enough to spot where a student had made his errors, they could be corrected. According to this teacher, the re-learning process took several times, because habits were hard to break. Mr. DiCarlo felt that if students could learn the habits right in the first place, they were so much easier to teach the next steps of the process. Mr. DiCarlo was soft-spoken and very serious with his students, and they respected him a great deal. He used hands-on activities, always learning new ways to teach a concept, changing and evolving, becoming better and better (Field Notes, 10/27/03). There was a certain level of frustration felt from this teacher over the state assessment system and the NCLB Act. He shared that classroom teachers should have the data they needed at the beginning of the year for current students, and he

wanted everything to be very organized. He knew that Oklahoma had plans to put a better assessment system into place, and he wanted that system in place "now." His students scored very high on the PASS CRTs, every year, with many students scoring a 100%, and all others in the 90th percentile range (Field Notes, 10/27/03).

Participant #5 (P5), Ms. Lynda Edwards

From observations of this teacher (Interview, 10/27/03), she appeared to have a great "passion" for teaching, working diligently to present her lessons in a very interesting way. Sharing that she and her children always liked the teachers who had made a class interesting, she felt this was her "key" to motivating students to learn. She wanted to be one of those teachers that students always remembered as making school "fun" at the same time they were learning. At the beginning of the year, Ms. Edwards was one of those teachers that students and parents requested because she did many creative hands-on activities to help students learn. Even with the lessons that required practice and drill, she worked to be lively and enthusiastic, impressing on the students a need to learn the skill so they would be ready for the next level. She also was quite concerned about what the NCLB Act could do to the teaching profession. She believed that many schools in Oklahoma were not ready and that many smaller schools could give up before they started (Field Notes, 10/27/03). Very concerned that all her students learn what they were supposed to learn, she pointed with pride to their achievement scores. Every year, 100% of her students scored above satisfactory

on the state test, with most of them scoring above the 90th percentile on the CRT. She searched for new and different ways she could teach her lessons, always asking for professional development activities and resources to make that happen (Field Notes, 10/27/03).

Participant #6 (P6), Mrs. Debbie Flores

Mrs. Debbie Flores (Interview, 11/07/03) was by far the most dedicated reading teacher, one who, every year, worked with all the students who were at least two or more grade levels below their peers in reading. They worked diligently to progress in reading two and sometimes three levels in one year. Many of her students were English language learners (ELLs), while others were behind because of various reasons, including low-motivation. Mrs. Flores taught with great enthusiasm, attempting to make classes enjoyable, and she had many students who respected her (Field Notes, 11/07/03). Even though many of her students began her class below grade level, every year 80% of her students were at or above the proficient level (70th percentile), with many of them scoring in the 90th percentile. In discussing her success, Mrs. Flores said that making the class "fun" was her first ingredient in getting students motivated. She also used incentives, both praise and literal rewards such as free passes or chewing gum, for everyone on Fridays or special holidays. She took her job very seriously as she strived to find the exact level of every one of her students at the beginning of the year, and from then on, she continued to benchmark their progress. Her charts

indicated the substantial growth of the students throughout the year. This teacher took the challenge personally when one of her students did not want to learn, seeing that as a formidable task but doing everything she could to change the situation (Field Notes, 11/07/03). Even with below-grade reading students, her reading scores were at the highest level in the district every year. Because of her success, she was the one of the most admired teachers in the entire school system. The district had asked her to train other teachers, especially in how to use the diagnostics in the software system. She had one of the hardest jobs in the district, and not many teachers could have handled this kind of pressure year after year (Field Notes, 11/07/03). The second important ingredient Mrs. Flores used in getting students to learn was making the lessons relative to the students' lives. This participant used extensive classroom data, graphs and charts, and involved students in the visual graphics, showing them how they had moved up or down on the bars and how well they performed (or not) in their daily work. She had discovered an important piece of the puzzle in education, showing her students where they were on the chart and how each one mattered. This appears to be a "disconnect" that was missing in many classrooms (Field Notes, 11/07/03).

Results of Interviews With Key Participants

Several themes emerged from the data collected in the interviews, with all of the issues central to the following research question: How had "highly effective" teachers found methods to use Oklahoma achievement test results and

other educational data to impact classroom instruction and improve learning that would ultimately result in "adequate yearly progress" for No Child Left Behind. Those themes are as follows.

Theme 1: The Oklahoma State PASS Standards

All six participants in this study (P1, P2, P3, P4, P5, P6) related that they used the Oklahoma PASS standards to plan their instruction, their daily lessons, and a timeline. They shared that their district had aligned its curriculum (K-12) within the last three years and had curriculum guides that had made the PASS standards much easier to follow. Mr. Callahan (Interview 3, 10/15/03) said his district was ahead of the time in comparison to other districts in the state, because not only had it aligned its curriculum, but the district had implemented a pre/post test process that gave more information than did the state achievement tests. Ms. Edwards (Interview 5, 10/27/03) shared that at the time the district had undergone curriculum alignment, the teachers were not aware of the impact that would have in preparation for NCLB. With alignment in all subjects, Ms. Edwards (Interview 5, 10/27/03) believed the district was a step ahead in being familiar with the PASS standards because the teachers had worked on the skills for so long. As highly effective teachers, they all agreed that they used the state standards to plan their lessons instead of relying on textbooks for a curriculum.

These teachers (P1, P2, P3, P4, P5, P6) also felt the state blueprint for each subject was helpful in knowing which of the PASS standards were going to

be tested so they could give depth to those identified skills. Mrs. Flores (Interview 6, 11/7/03) said the district's three-year effort in curriculum alignment was helpful, but that the test blueprint for the PASS test was most important. Those skills were outlined into categories; however, she said some of those categories were vague. Mrs. Arbello (Interview 1, 10/3/03) said third grade teachers at her site did use the list of skills included with the NRT results, with those skills listed in categories with the number of items tested in each area. They used that report as a "blueprint" of the skills tested. The problem that the third grade teachers expressed was that the NRT test standards did not highly correlate with the PASS standards. Mr. Callahan (Interview 3, 10/15/03) said it was "tough" to teach the PASS standards and teach to the SAT9 (the NRT), but he also used the "blueprint" or list of skills that came with those reports. Mrs. Blair (Interview 2, 10/6/03) said she did teach to the PASS objectives, but those changed from year to year; so, she also used the SAT9 "blueprint" in addition. These teachers used a combination of both standards and blueprints in planning instruction, with the fifth grade teachers using the blueprints more and the third grade teachers relying more on the standards.

In addition, these teachers (P1, P2, P3, P4, P5, P6) said they documented when and how they taught their PASS standards. Mr. Callahan (Interview 3, 10/15/03) shared his method of tracking the skills along with the dates they had been taught (Appendix B). Ms. Edwards (Interview 5, 10/27/03) said she used a

cross-reference method with her textbook, marking each of the PASS standards and those identified from the blueprint, so that the most important skills would be taught in depth. Mr. DiCarlo (Interview 4, 10/27/03) said he dated the skills when they were taught, but that he could not just teach to those standards because there were "so many more things that students need to learn."

These six teachers (P1, P2, P3, P4, P5, P6) said they had the PASS guidelines, but they also had many years of experience at or near the same grade level. They said tests changed, test companies changed, and guidelines changed. They also said the PASS standards changed periodically. Mr. DiCarlo (Interview 4, 10/27/03) knew what his students were supposed to learn for the fifth grade; and knowing what to teach was not the problem; his concern was that there were far too many skills to effectively teach them all in depth. Throughout the interviews, all other participants echoed his sentiments.

The participants all agreed (P1, P2, P3, P4, P5, P6) that they must align their instruction to the standards and to the blueprints, but they were concerned about the time that it took to teach every item in depth. Mr. Callahan (Interview 3, 10/15/03) said skills needed to be streamlined so students and teachers both were not so stressed. Both Mrs. Arbello (Interview 1, 10/3/03) and Mrs. Blair (Interview 2, 10/6/03) said they had no choice but to attempt to teach the standards and the test items, so that students could perform their best on the test. Mr. Callahan (Interview 3, 10/15/03), Mrs. Arbello (Interview 1, 10/3/03), and

Mrs. Blair (Interview 2, 10/6/03) said they had their skills and test items matched with a timeline. Mrs. Arbello (Interview 1, 10/3/03) said she could get all the skills taught, but perhaps not in the same depth as others; with the important ones (Interview 1, 10/3/03). All teachers (P1, P2, P3, P4, P5, P6) said they could be much more effective in the classroom if they had fewer skills, and if fewer skills were tested. Mrs. Blair (Interview 2, 10/6/03) said it was extremely difficult to satisfactorily teach all PASS standards to ALL children, adding that she did not know how any teacher could accomplish that. These teachers (P1, P2, P3, P4, P4, P6) taught hurriedly every year, and they shared that teaching was not very enjoyable because of the time involved, causing enormous stress.

These participants (P1, P2, P3, P4, P5, P6) did not know of any elementary teachers who failed to follow the standards, especially in reading and math, and they felt their instruction was aligned to the standards. They felt the skills included in the standards were so ingrained that they did not know any other method to teach. In addition, they felt that while some children may do well, other children would not and that while teachers could introduce some concepts, not all could be mastered. At one time, teachers had simply taught what was outlined in the textbooks, starting in chapter one and covering as much of the text as possible. These participants (P1, P2, P3, P4, P5, P6) did not believe that many teachers in the district still did that. Mrs. Flores (Interview 6, 11/7/03) had enjoyed teaching when she could include "fun" activities, but now she stayed with

the standards. All participants (P1, P2, P3, P4, P4, P6) learned they could not cover the entire text, and they designated what was most important. The state of Oklahoma, with help of state teachers, had designated, and all felt that was best for students.

Table 8: Matrix of Findings on Oklahoma State Standards

Theme 1: Findings and Sources of Data Triangulation Sources of Data: I = Interviews, S = Surveys, O = Observation, D = Documents	I	S	O	D
1. Teachers planned instruction according to Oklahoma PASS standards.	x	x	x	x
2. Teachers also aligned their instruction according to test blueprints.	x	x	x	x
3. Teachers documented the teaching of the PASS standards.	x	x	x	x
4. Teachers thought time did not allow them to teach all PASS standards.	x	x	x	
5. Teachers had established good instructional practices with standards.	x	x	x	
6. Teachers felt that a fewer number of standards would reduce stress.	x		x	
7. Effective teachers believed they would be more effective as teachers with fewer standards and fewer standards tested.	x		x	

Notes: Document/Artifact for #1 = Appendix B
Document/Artifact for #2 = Appendix C
Document/Artifact for #3 = Appendix D

Theme 2: Oklahoma Assessments and Blueprints

With the present assessments, Oklahoma school districts must show Adequate Yearly Progress (AYP) for the No Child Left Behind Act (NCLB) with all groups taking the same test at the same grade level. AYP would then be decided by taking the scores of children from last year and comparing them to the children's scores in the ensuing year. All children's and subgroups' scores would be reported, including special education and English language learners (ELLs). If students did not make AYP for two years in a row in their subgroup, sanctions would be imposed by the state and federal governments, including the withholding of funds. Participants in this study (P1, P2, P3, P4, P5, P6) did not believe the assessment process fair to all children.

Disagreeing with the mandates to use the same test for all children at the same grade level, the six teachers interviewed (P1, P2, P3, P4, P4, P6) did not want last year's data on their students but wanted current data on the students at the beginning of the year. Mrs. Blair (Interview 2, 10/6/03) did not agree with all students taking the same test because they were not at the same point at the beginning or at the end of the year. She especially disagreed giving the same test to her special education and ELL children. Acknowledging that there were test accommodations, she emphasized that every student in her class would take the same test, no matter what their grade level performance. What Mrs. Blair found even more frustrating was not having current data on her current year's students,

believing that comparing one year to another was comparing "apples to oranges" (Interview 2, 10/6/03). Believing that teachers should raise each child's scores from his/her level at the beginning of the year, Mrs. Arbello (Interview 1, 10/3/03) questioned why schools could not use the actual progress a student had made from year to another. Mrs. Arbello said, "Why do we all have to be at the same place at the same time? When that does not happen in the real world?" (Interview 1, 10/3/03). Ms. Edwards (Interview 5, 10/27/03) agreed, believing the state assessment data were not really helpful, because it was from last year's set of students, requiring teachers to do their own assessments and diagnostic tests at the beginning of the year for current students. Mr. DiCarlo (Interview 4, 10/27/03) stated that it was too late to "fix what I have messed up" or did not teach to the previous year's students. He added that while teachers could spot gaps or holes in their instruction, the students each year were a different group at different levels with different needs. Mr. DiCarlo stated, "Where last year's students were strong, the new group may be weak" (Interview 4, 10/27/03). He added that teachers should pre-test their current students to see where they were performing. Mrs. Flores (Interview 6, 11/7/03) said she always looked at students' test results from the previous year, but she said that did not tell her where her current students would be. Mrs. Flores stated, "The state says it is tracking the teacher's and the school's instruction, but it is so much more than that" (Interview 6, 11/7/03). She asked how the state could compare one year's

students to the next year's group because one class may be really bright and another quite average. Mrs. Flores asked, "Where does that leave us when we get a class that is less than average?"(Interview 6, 11/7/03). Mr. Callahan (Interview 3, 10/15/03) agreed with all the other teachers interviewed. He believed the data the school had each year needed to go on to the next grade level to the teacher who would have those students, so that each year, a teacher would have data on his/her current students and not on last year's. These teachers (P1, P2, P3, P4, P5, P6), ones who were leaders in their school community, seemed to have a deep understanding of the state assessment process and what would and would not show gains or Adequate Yearly Progress.

Some of the teachers found the methods of reporting the assessment data difficult to read and understand. Mrs. Blair (Interview 2, 10/6/03) believed the categories and subgroups of standards on the group reports to be very "vague." Mrs. Blair said teachers had rather the categories be more well-defined so that they would not have to guess what a category such as "textual reading" meant. Mr. Callahan (Interview 3, 10/15/03) said that some of the areas tested were very hard to understand, such as "appreciates reading" or "problem solving." He asked that the state make very clear the skills and areas tested so teachers could better plan their instruction. Ms. Edwards (Interview 5, 10/27/03) said that teachers did not want to know the exact questions on the test, but they would like more information on the categories and subgroups because the reports were so hard to

read and understand. Mrs. Flores (Interview 6, 11/7/03) agreed, adding that the test information was also vague, with a category such as "literary terms" much too broad. Mrs. Flores added that the blueprint did not give enough information, asking that categories on all test blueprints be more specific. Asking why the state could not share exactly what would be tested, Mrs. Flores (Interview 6, 11/07/03) said that at the most, the blueprint was "guessing in the dark." These teachers asked for data reports that would be well understood along with well-defined skills on blueprints.

Because the state test data were not easy to read, nor was the data on current year's students, these teachers (P1, P2, P3, P4, P5, P6) were forced to create their own diagnostics and pre tests. Using his own unit and pre tests, Mr. DiCarlo (Interview 4, 10/27/03) said that he did not use state test data to plan instruction or interventions; he used his own assessments and worked one on one with students daily to see what should be re-taught. Believing that teacher-made tests gave more information than the state test data, Ms. Edwards (Interview 5, 10/27/03) thought the district's pre tests that were skill-specific were much more helpful when identifying "low" areas needing improvement. Ms. Edwards added that she made most of her own assessments, looking for material that would aid in teaching a concept from a different angle, hoping for mastery from all her students. Preparing more than her own teacher-made assessments, Mrs. Flores (Interview 6, 11/07/03) said that she did her own diagnostic tests with the aid of

computer software. Mrs. Flores believed that diagnostics should be an integral part of instruction and were so much helpful than the state assessment data.

When asked how they did or could use the state assessment data, the six participants (P1, P2, P3, P4, P5, P6) gave various answers. At least four of the participants mentioned the following statements.

- At the end of school, when test scores arrive, teachers looked to see what skills were low or lacking. (P1, P2, P3, P4, P5, P6)
- Teachers only looked at the state test results once or twice after they got them. (P2, P3, P4, P5)
- Teachers used their own assessments and their own experiences to track student progress and to plan for AYP and their own data from were more key to planning instruction than state data. (P1, P2, P3, P4, P5, P6)
- Teachers pre tested current year's students rather than using the state data from a previous year. (P1, P2, P5, P6)
- State data on current year's students would be more helpful. Teachers had to create their own pre tests, skills tests, and other assessments. (P1, P2, P3, P4, P5, P6)
- By looking at norm-referenced test data, teachers believed they were working in two directions because the state standards did not highly correlate with the NRT. (P1, P2, P3, P6)

- Teachers did not use state data to make charts, graphs, or timelines. (P2, P3, P4, P5)
- Assessment should be a part of everyday instruction, not a once-a-year item. (P1, P2, P3, P4, P5, P6)

One teacher, Mrs. Flores (Interview 6, 11/7/03), said that she did use the state achievement scores throughout the year to see what areas were weakest and where she needed to work harder. She conceded that those test results were all that she had to plan for her next year's instruction, adding that she did not know if using the state data were the best way but it was the only way she had. Mrs. Flores added that she did give placement tests to students at the beginning of the year and what would "really be nice" would be to know where her current year's student were when she first got them (Interview 6, 11/7/03). She also added that teachers were not supposed to know what was on the test, but that students would make comments about what they had or had not studied. Mrs. Flores asked, "Why is the test material such a deep, dark secret that none of us are supposed to know?" (Interview 6, 11/7/03). Mrs. Flores and other participants believed that they did not have enough information from the state data for it to be truly useful.

Additional comments from teachers about the state assessments included those about special education students. These teachers (P1, P2, P3, P4, P5, P6) concurred that special needs students should be mainstreamed if at all possible, and they did not mind having them in their classrooms. However, they thought

the state was doing a disservice to those students by making them take a state test that could possibly be several grade levels above their abilities. The participants (P1, P2, P3, P4, P5, P6) were knowledgeable about the "safe harbor" clause in the state AYP requirements where Oklahoma special education and ELL students were compared to last year's students and did not have to meet a specific score on the API. In addition, these safe harbor students could not score less than 10% from the score of last year's students (OSDE, Student Assessment, 2003). While the state and federal governments required Individual Education Plans (IEPs) to be made for special education students with modifications for grade level work, these students for the Oklahoma API and NCLB were being tested at the same level as all other students. For instance, if an eighth grade student was performing at the third grade level, he/she must take the eighth grade test. The state allowed for 1% of all Oklahoma students to be tested with an alternative test (OSDE, Student Assessment, 2003), and the necessary portfolios entailed many hours of work for each student. These six "highly effective" teachers (P1, P2, P3, P4, P5, P6) believed that 1% would not be nearly enough to provide a "safe harbor" for Oklahoma's special education or ELL students. These teachers were responsible for making modifications for IEP students in their classrooms, while at the same time, they administered a test far above students' performance levels. They asked for tests to be administered at the students' ability levels, because that range could be several grade levels below and students would, in turn, be frustrated.

Table 9: Matrix of Findings for Oklahoma State Assessments and Blueprints

Theme 2: Findings and Sources of Data Triangulation Sources of Data: I = Interviews, S = Surveys, O = Observation, D = Documents	I	S	O	D
1. ALL students at same grade level should not have to take SAME test.	x		x	
2. State data on current students would be more helpful than state data on last year's students.	x	x	x	
3. State data should be more user-friendly, easier to read and understand.	x		x	x
4. Test results and blueprints, categories and sub-areas, should be well defined.	x		x	x
5. State standards at third grade did not highly correlate with the NRT.	x	x	x	
6. State test data were a cursory part of teachers' planning and instruction.	x	x	x	
7. Teachers used their own assessments and classroom data to plan for Adequate Yearly Progress (AYP).	x	x	x	x
8. Teacher-made assessments were more helpful than state data to drive instruction.	x	x	x	x
9. Teacher-made data and assessments were an integral part of everyday instruction.	x	x	x	x

Notes: Document/Artifact for #3 = Appendix D

Document/Artifact for #4 = Appendix C

Document/Artifact for #7 = Appendix E

Document/Artifact for #8 = Appendix F

Document/Artifact for #9 = Appendices E & F

Theme 3: Teachers' Use of Educational Data

The literature remained clear on using data to drive and improve instruction and as a reflective educational tool, and the six teachers interviewed (P1, P2, P3, P4, P5, P6) all accepted that conclusion. Ms. Edwards (Interview 5, 10/27/03) said that at one time very few teachers knew how to use or disaggregate the data, so few knew which data were helpful and which were not. Thinking the use of data were very important, Ms. Edwards said that she had asked the special education teachers to help her interpret the results. She said the curriculum director had trained them how to find the percentiles, by group and individually, and compare them to the previous year. That director had also trained teachers in the difference in the norm-referenced and criterion-referenced tests. At the beginning of the year the curriculum director had brought copies of the data to all sites and explained what the different data meant, but when the teachers used the data only once or twice a year, Ms. Edwards (Interview 5, 10/27/03) said the explanations would not be remembered. Mrs. Flores (Interview 6, 10/7/03) asked about Adequate Yearly Progress (AYP) and improving every year until the district "hits the magic number" in the year 2013-14. She wanted to know where the district and state went from there, stating "Did teachers keep on having to improve? What would be enough?" (Interview 6, 10/7/03). These teachers wanted to become "data-driven," but several thought only the state achievement test data were "real" data.

Even though all participants (P1, P2, P3, P4, P5, P6) agreed using state and classroom data were important, they disagreed on which data were useful and helpful. Because the state data were the only standardized data she had, Mrs. Flores (Interview 6, 10/7/03) said she used state data in conjunction with her own assessments. She said, "I use my own data, because I may keep on teaching without knowing whether students have learned what they are supposed to have learned" (Interview 6, 10/7/03). Diagnostic tests and placement assessments had become very important for Mrs. Flores, a very real part of her instruction. Of the two types, Mrs. Flores believed her own teacher-made assessments were much more useful than the state data (Interview 6, 10/7/03).

Just as others had said, Mrs. Blair (Interview 2, 10/6/03) concluded the state data were only a "guideline" for her and others in planning and interventions, explaining that they used pre tests as a "blueprint" to know what interventions were needed. Mrs. Blair used many "hands-on" activities, such as chalkboard demonstrations and oral reviews, as data results. She added if the state wanted her to prove progress, then it should have her test her students at the beginning of the year and then again at the end of the year to show real AYP. Mrs. Blair said that last year's state data told her if her students were low in spelling or in reading comprehension, but those students had moved on to the next grade. She said that some years, the results looked like she had done an "awful" job of teaching, but that she may have made more than a year's progress with

some students; however, the test results would not show the progress she had made (Interview 2, 10/6/03). For Mrs. Blair, the state data meant very little when the results showed she was not a very good teacher while in reality, she may have done an outstanding job with the children she had taught.

Ms. Edwards (Interview 5, 10/27/03) agreed that classroom assessment were more accurate than the state data, with one possible exception. Ms. Edwards explained that if a student did really well in class but not on the test, the child might not be working as independently as he should, with parents or teachers helping the child too much with his work (Interview 5, 10/27/03). While all teachers agreed that teacher and student should be responsible for a student's learning, they argued that using the state data were not the way to accomplish accountability. All six teachers (P1, P2, P3, P4, P5, P6) agreed that their own classroom data were much more beneficial in showing actual progress.

All participants (P1, P2, P3, P4, P5, P6) agreed that data on the current year's students would be the most helpful. Creating her own assessments for her current students with pre tests, Mrs. Blair (Interview 2, 10/6/03) administered skills tests, along with other assessments such as computer-assisted software. As a group of teachers, Mrs. Blair stated that they would rather see data on the second grade students who were becoming third graders, but the state did not furnish data for that group. Mr. Callahan (Interview 3, 10/15/03) said his grade level colleagues used the state test scores only at the beginning of the year to see

what areas were low so more emphasis could be put on those. "Again," Mr. Callahan said, "we need the state data that goes with the children that will have them next year." (Interview 3, 10/15/03). Mr. Callahan said that by the time teachers in his group could meet in grade-level meetings to disaggregate the data, they all had new students.

All six teachers (P1, P2, P3, P4, P5, P6) thought they needed data on the current year's students in order to prepare for NCLB and AYP. Teachers knew that Oklahoma had announced plans to implement CRT assessments in reading and math from third through eighth grades; however, these teachers would still have to create their own activities and data to benchmark progress for their students. Mr. DiCarlo (Interview 4, 10/27/03) said he benchmarked student progress in math by giving them an assessment every five lessons, tests that "looped" or reviewed previously taught skills. This teacher had learned to use his own data to re-teach and plan further instruction. Mr. DiCarlo said he knew his students had learned what they were supposed to have learned when they made an 80% or better on classroom assessments. If they made below that, the teacher worked with the student individually to see where the problem lay (Interview 4, 10/27/03).

Mrs. Arbello (Interview 1, 10/3/03) said her grade level had created an item analysis from teacher-made pre tests, placing more emphasis in certain skills in reading and math. Stating that her group was doing everything that it could,

Mrs. Arbello said they were getting practice tests ready so that students would be prepared in April for the Oklahoma tests. Mrs. Blair (Interview 2, 10/6/03) said her group used all the classroom data they could find, making charts using the district pre test data. That chart, and list, became a blueprint or guide for her in knowing what areas needed more instruction. Mrs. Blair said teachers knew how many skills the students had missed from the categories, and with the percentages, they knew which skills needed most work. Explaining that the district teachers were making a test bank for identified "low" skills, Mrs. Blair thought they would benefit from all other teachers expertise in those areas. "The people in the 'trenches' are the ones we can learn most from," Mrs. Blair said (Interview 2, 10/6/03). She added that district pre tests had helped her class immensely because those results from a computer item analysis were on grade level and broken down into specific groups.

Both Ms. Edwards (Interview 5, 10/27/03) and Mr. Callahan (Interview 3, 10/15/03) agreed with Mrs. Blair (Interview 2, 10/6/03) about the value of the district's pre tests and the data that the computer software could give them on individual students. Agreeing that the district pre tests were valuable, Mrs. Flores (Interview 6, 11/7/03) said that was not enough. She tracked individual and class progress with charts and graphs to show her students how well (or not) they were performing. Mrs. Flores said the graph helped her as a teacher, and it helped the students to see how much work they would have to accomplish in order to

perform at grade level. She said if the whole class worked at an average level, they could see that on their chart; however, she expected an 80% or better, using rewards and incentives. Mrs. Flores (Interview 6, 11/7/03) said if her students needed more work on certain skills, she taught the lessons in a different way, with more hands-on activities, which in turn took more time. Mrs. Flores also kept a folder on each student, data for herself as a teacher, for the student, and for the parent. Because they did not want to rely solely on state data, these highly effective teachers had found very creative ways to use their own classroom data.

Table 10: Matrix of Findings for Teachers' Use of Educational Data

Theme 3: Findings and Sources of Data Triangulation - Sources of Data: I = Interviews, S = Surveys, O = Observation, D = Documents	I	S	O	D
1. Using data to drive and improve instruction and as a reflective educational tool was an important part of the teacher's instructional process.	x	x	x	
2. Teachers used state data for some information, but only as a "guideline."	x	x	x	
3. Use of data from classroom assessments on the current year's students was the most beneficial to teachers in showing students' actual progress.	x		x	x
4. Teachers created their own assessments to show progress and benchmark student growth for NCLB and AYP.	x	x	x	x
5. Diagnostic and placement assessments were an integral part of classroom instruction.	x	x	x	x

Notes: Document/Artifact for #3 = Appendix F
Document/Artifact for #4 = Appendix G
Document/Artifact for #5 = Appendices E & F

Theme 4: Proven Practices/Interventions

The six "highly effective" teachers appeared well trained in the use of instructional materials and in effective teaching/learning methods. These teachers were leaders in their buildings, modeling the practices and interventions that made a difference in children's lives and created varied learning opportunities for their students. From the descriptions of the participants, each one had different elements that proved to be strengths for motivating students and in getting them to learn. The proven practices of these six teachers were also varied and different, but several categories emerged as common practices and effective teaching methods. The practices or interventions that all six participants implemented were categorized as follows.

1. Teachers made their own item analyses from district or teacher-made pre tests. These provided blueprints or lists for teachers to follow throughout the year. They made timelines from these item analyses and rank ordered the skills, starting their teaching with the weakest skills. Two did the item analysis by hand (P1, P2), while other participants used the item analysis that was computer generated from the district pre test. (P3, P4, P5, P6)
2. Teacher-made assessments or placement tests provided the most helpful information to all teachers at the beginning of the year. Teachers tested students in various ways, giving these at the beginning of the year, and throughout both semesters to benchmark student progress. They said it

helped them to understand at what level their children were performing and which students needed the most help. Teachers had also made a district pre/post test that they administered at the beginning and end of every year to track progress and gains. (P1, P2, P3, P4, P5, P6)

3. Individual one-on-one tutoring proved very useful for students who had problems learning the material. Teachers found that while all concepts could be introduced, not all could be mastered. If there were not an after-school program, many teachers stayed after on their own time, working with students until the concepts were mastered. (P1, P2, P3, P4, P5, P6)
4. Cooperative or group work proved a prominent teaching method in these teachers' classrooms. This work ranged from using verbal tests, giving oral reports, finding examples, playing math games, doing chalkboard activities, and doing oral review or "looping." Group work gave stronger students an opportunity to help another student, and peer teaching provided help for students were having problems. (P1, P2, P3, P4, P5, P6)
5. Demonstrations and "hands-on" activities by the students were very helpful in the learning process. Students liked the "hands-on" activities and many learned better that way. Teachers felt that students were different types of learners today because of television and video games, believing that teachers could not instruct the way they always had and get any better results. Manipulatives, such as rulers to measure and pattern

blocks to build, allowed students to "see, feel, and touch." The problem teachers found was not having enough hours in their day to do the hands-on activities because they took more time. One participant (P5) said her class demonstrated in "some way, every day." (P1, P2, P3, P4, P5, P6)

6. Guided practice, reading and math exercises with the teacher's help, was a prominent method used in the effective teachers' classrooms. This guided exercise was with worksheets and textbook assignments. These teachers watched their students work so they could give immediate feedback and interventions. Most teachers closely watched their students solve problems and those needing help would get one-on-one assistance. These teachers were "masters" at helping students find where they had made their errors and correcting the problem-solving process. (P1, P2, P3, P4, P5, P6)
7. "Looping" or cumulative reviews was a frequent practice in these teachers' classrooms, often in the form of an oral review. These teachers believed that once a student had learned a skill, that skill should be reinforced and used often during the year so that students would retain the information in their long-term memories. (P1, P2, P3, P4, P5, P6)
8. To provide visual learning pictures, graphs, charts, and timelines were often used as instructional tools. Students could see where they had performed by looking at a chart or graph, tracking growth or gains throughout the year. This showed the parents where students were strong

or weak. Teachers used data from many sources, including student feedback, with both formal and informal data. (P1, P2, P3, P4, P5, P6)

9. Skills Tests and Practice tests were often used, with students learning the format of standardized achievement tests. These teachers felt that in order to prepare students properly for the state tests, the students should be taught "test skills"; however, teachers found giving the practice tests time-consuming. (P1, P2, P3, P4, P5, P6)
10. Computer software was used to supplement classroom materials, for remediation and for diagnostics. According to these teachers, this information facilitated the teaching/learning process more than state assessment data did. The classes were on a rotating basis in using the computer labs, so time limited the use of software. (P1, P2, P3, P4, P5, P6)
11. Portfolios or folders were kept on all students, information for teachers, students, and parent. This also provided documentation for students who needed other interventions, such as gifted, special education, or English as a second language classes. (P1, P2, P3, P4, P5, P6)
12. Other resources were frequently used to supplement the teachers' materials, including lower-level materials, task cards, manipulatives, games, computer software, old encyclopedias, and the Internet. (P1, P2, P3, P4, P5, P6)

Table 11: Matrix of Findings for Teachers' Proven Practices and Interventions

Theme 4: Findings and Sources of Data Triangulation Sources of Data: I = Interviews, S = Surveys, O = Observation, D = Documents	I	S	O	D
1. Item analyses from district or teacher-made pre tests were used.	x	x	x	x
2. Teacher-made assessments gave valuable information for placement.	x	x	x	x
3. Individual tutoring proved important for many students.	x		x	
4. Cooperative work afforded group learning and review.	x		x	
5. Demonstrations and "hands-on" activities aided the learning process.	x		x	
6. Guided practice proved a prominent teaching method used.	x		x	
7. "Looping" or cumulative reviews were frequent practice.	x		x	
8. Graphs, charts, and timelines supplied instructional tools.	x		x	x
9. Practice showed students the format of standardized achievement tests.	x		x	
10. Computer software supplemented classroom materials	x		x	x
11. Teachers kept portfolios or folders on all students	x		x	
12. Other resources supplemented the teachers' materials	x		x	

Notes: Document/Artifact for #1 = Appendix H

Document/Artifact for #2 = Appendices E, F & I

Document/Artifact for #8 = Appendix J

Document/Artifact for #10 = Appendix K

Theme 5: The Issue of Time

The issue of "time" surfaced over and over during the interviews of the six participants (P1, P2, P3, P4, P5, P6). Extra "time" had made it possible for these "highly effective" teachers to collaborate with others on instruction, to extend their school days above and beyond others; to find new, innovative ways to keep students interested, and to keep themselves updated on the state standards and assessments. They all (P1, P2, P3, P4, P5, P6) concurred that there was never enough time in the school day to plan instruction, grade papers, and raise student achievement. Working long hours before and after school and during the summer was a common practice for those interviewed. The willingness to give and apply the time to study the standards and blueprints, to pre test students, and to make item analyses may have been the most significant value that separated the effective teachers from those not as successful.

To these effective teachers (P1, P2, P3, P4, P5, P6), the constant interruptions of class time proved to be their toughest problem or issue to solve. They wanted to preserve as much instructional time as they could for their students. At the elementary level, the interruptions came in the form of pullout programs such as those for the gifted, talented, speech, remedial, music, physical education, and special education students. Teachers found planning a large block of time for reading or math very difficult.

In addition, the teachers (P1, P2, P3, P4, P5, P6) did not have time to collaborate with each other except before or after school. Time was the one problem that these teachers specifically said they needed help solving from their principals or from other school leaders in scheduling their school day. Individual comments from this group varied.

Stating that teaching took more "time" than any of them had ever dreamed, Mrs. Flores (Interview 6, 11/7/03) said that she and her grade-level group could not get all the skills taught by test time. The first year, Mrs. Flores said she played "catch-up" in reading all year, getting the students ready for test time. Now, Mrs. Flores said she spent all her time teaching the standards by the end of the first semester. Then, she worked on reviews, depth, and mastery during the second semester, adding that the method worked for her and her students. Her class worked on numerous practice tests and other test preparation, especially terminology. Mrs. Flores said her planning period was never enough to prepare for the next week's instruction or grade papers, much less spend time analyzing data (Interview 6, 11/7/03).

To make sure that all her students learned so much in a short amount of time, Mrs. Blair (Interview 2, 10/6/03) said she had time for little else. She also started from the very first day of school and worked through both semesters getting everything taught. Mrs. Blair said the "push" started just before test time in order to get everything "covered." She said there was never enough time, and

that she did not want her class to be that way, but she felt she had no choice. In addition to teaching duties, Mrs. Blair said she has after-school duties, such as bus monitoring or grade-level meetings. She said she used her summers as wisely as she could, looking for new methods or new manipulatives for hands-on activities. She said she was always looking for new ways to teach her children as well as new ways to use the data she had on each child (Interview 2, 10/6/03).

Stating that she started from the very first day of school, Mrs. Arbello (Interview 1, 10/3/03) said her students worked hard every day until test time. She and her students went through each of the skills one by one with plenty of "practice, practice, practice." After test time, Mrs. Arbello's class relaxed and did more hands-on and group activities, such as practicing long division in groups because some students needed the help of others. A very dedicated teacher, Mrs. Mrs. Arbello said she went to work at 6:30 a.m. every morning and seldom left her classroom before 6:00 at night. She found coming early and staying late the only way she could manage the preparation for state standards and assessments (Interview 1, 10/3/03).

From the Oklahoma State Department of Education (OSDE, Student Assessment, 2004), Grade five PASS standards in Language Arts included four sets of state standards. There were 127 skills and objectives listed under the standards just for language arts; from the first set there were 56 skills, from the second set, 53; from the third set, 11, and from the fourth set, 7.

After giving his pre tests, Mr. DiCarlo (Interview 4, 10/27/03) said he could often omit the first few lessons in a core curriculum after looking at the test results. He did not have time to look at the pre test data during the school day or week, but he spent his weekends, usually Sunday night, looking at school work and "gearing" up for the week ahead. Mr. DiCarlo said the first year he taught, he did not get all the skills taught by test time, but that his class "crammed" the last two weeks before the test. After that year, Mr. DiCarlo gave pre tests that would allow him to leave out review lessons that students did not need. He said if his students had retained material from the previous year, he had less work to do; however, if they had not retained the material, he had to re-teach or review more at the beginning of the year (Interview 4, 10/27/03). Throughout the year, Mr. DiCarlo continued to review or "loop" the skills to make sure students retained the material by test time.

Referring to her fifth grade level, Ms. Edwards (Interview 5, 10/27/03) said that teachers at her site never had time to teach all the skills or to analyze the data. She said the teachers had little time to plan individually during school and even less time to plan as a group. Her group looked at the state data at the end of the year and then again at the beginning of the year, but that was all the time they had. She did say the computer software and placement assessments her group had prepared were valuable and time-saving. She had commented that technology was "great" because now they could design a test, save it, and then have only

minor adjustments before they could use the material for the next year. With the state tests scheduled for April and the results reported by September, Ms. Edwards did not see how the state results could assist them. School started in the middle of August, so the state test results would not be helpful nor would not save them any time. They would still have to do their own placement tests and item analyses. Ms. Edwards said that by September, they would have their year already planned (Interview 5, 10/27/03).

Emphasizing that teachers worked Saturdays, weekends, and after school in instructional preparation, Mr. Callahan (Interview 3, 10/15/03) said it was difficult to get everything done in a school day. Besides instructions, teachers faced paperwork in the form of grading along with local, district, and state reports that had deadlines. In addition to after school duties, Mr. Callahan could not "adequately" teach all the skills by the first of April and test time, adding that he taught "fast and furious" and taught more lessons each week than what was considered normal for that grade level. Mr. Callahan came close to teaching all the skills in math, but not in language arts. In English, he chose the important skills and went back after test time to fill in the "gaps." Having no planning time during the day and often after-school duties, he felt "overwhelmed." Mr. Callahan believed that the skills needed to be streamlined so the teachers and the students would not be so "stressed" (Interview 3, 10/15/03).

Table 12: Matrix of Findings for The Issue of "Time" for Teachers

Theme 5: Findings and Sources of Data Triangulation Sources of Data: I = Interviews, S = Surveys, O = Observation, D = Documents	I	S	O	D
1. Teachers did not have enough "time" during the school day for teaching/learning.	x		x	
2. Teachers could not "adequately" teach all the standards by the first of April.	x		x	x
3. Constant interruptions of class time took away from learning "time."	x		x	
4. After-school duties and meetings took additional "time" from preparation.	x		x	x

Notes: Document/Artifact for #2 = # of Language Arts Standards, Page 136
Document/Artifact for #4 = Calendar, Appendix L

Theme 6: Discipline and Parental Involvement

These six effective teachers understood that within their classrooms children should respect each other and their teacher, and they all (P1, P2, P3, P4, P5, P6) networked a process where parents and educators worked together. They felt that discipline was a major key in the teaching/learning process, and they had made an intense effort to involve parents in the discipline procedures, especially self-control. Each of the participants expressed a concern that good classroom discipline had eroded in the past years, and each wanted parents to take ownership in helping to develop order and respect that students must have for the learning process. As in all the other issues presented for these participants, each of these

"highly effective" teachers (P1, P2, P3, P4, P5, P6) created a good discipline structure within their classrooms in a very different way.

Being highly attentive and perceptive to his students' needs was Mr. DiCarlo's way of establishing a classroom climate where discipline was not a problem. Mr. DiCarlo (Interview 4, 10/27/03) gave his full attention to his students every day, listening to words, watching body language, and tuning in to any problems the students may have had. He seemed to know or perhaps felt when a student could not solve a problem or was frustrated. This value helped him to be highly effective and to have good discipline. To teach well, he could not have constant interruptions, because that distracted him and diverted the attention of other students. Mr. DiCarlo thought that discipline was becoming a bigger and bigger problem every year, believing that it would continue. He said parents no longer supported the school in consequences for bad behavior, and he felt the parents' non-supportive attitudes, added to the students' bad behavior, was more than schools should have to manage (Interview 4, 10/27/03).

With a commanding voice, Mr. Callahan's impressive manner kept his students in their seats and paying attention, and his presence was respected by the students. Mr. Callahan (Interview 3, 10/15/03) thought discipline was very important to the learning process, and that no matter how bright the students were, he could not teach them if all he dealt with was behavioral problems. The time in handling discipline slowed Mr. Callahan down, and every year, he said he had

more and more discipline problems with less and less teaching. If communications between parents and teachers were good, he believed the students were better behaved; however, the teachers could not always get the parents involved. "Sometimes," Mr. Callahan said, "we have as much teaching to do with parents as we do students" (Interview 3, 10/15/03).

Relying on creative ways to make students behave and keep them interested was Mrs. Blair's method of establishing good discipline from the very first day of school. Mrs. Blair (Interview 2, 10/6/03) said she spent several days at the beginning of the year gaining control of the children and establishing discipline for a learning environment. She believed discipline had the biggest impact on student learning, and she called parents often to get them involved. If she could not gain the parents' support, she felt it was her job to teach self-control and self-discipline. Mrs. Blair said one child could distract the whole class with valuable learning time lost. She felt that children of the last decade no longer respected their elders, with television and video games a major impact on home life that carried over into school. Mrs. Blair thought today's children were afraid of nothing and attempted to be "cool" at school, imitating the models they saw on television and the movies (Interview 2, 10/6/03).

Keeping her students interested and involved was Ms. Edwards' method of avoiding discipline problems in the classroom. Ms. Edwards (Interview 5, 10/27/03) felt if students were interested in the subject matter and they were busy

from the moment they stepped into the classroom, that students would exhibit good behavior. She felt that discipline and a student's respect, for his teacher and classmates, was the number one element in the teaching/learning process. What Ms. Edwards wanted to do was to keep students interested in what was taught, so she searched for ways to make even the most mundane lessons a little livelier. She wanted students to love learning. Ms. Edwards thought if all teachers would do that, they would have fewer discipline problems. In addition, Ms. Edwards said students inherently knew when another students' behavior was not appropriate, and those students had ways of displaying peer pressure and their influence to stop the bad conduct. Ms. Edwards also felt that teachers should be flexible with students and not so rigid in their rules (Interview 5, 10/27/03).

Completely agreeing that discipline impacted teaching and learning, Mrs. Arbello (Interview 1, 10/3/03) worked to build a climate where students were emotionally secure. She taught her students that she loved them from the very first day of school. She thought it imperative that a child have the emotional security with his teacher in order to learn well, and she worked diligently the first week of school to convince each and every child that she loved them. "If they know I love them, they will do a better job," Mrs. Arbello stated (Interview 1, 10/3/03). Also, she felt it her duty to teach them respect and responsibility, caring about what they accomplished. If she could bring about the right climate in her

classroom, those students would do their best in everything they did, "for now and for a lifetime – even into future jobs they would hold" (Interview 1, 10/3/03) .

Stating that all classes had different personalities and problems, Mrs. Flores (Interview 6, 11/7/03) related the lessons she taught to the real world in the most interesting way she could present them. Some of her classes were cordial and friendly, while others were not, and that set a certain mood in her classroom, for her and for the other students. Relating all her lessons to the real world and getting students to understand how a concept or skill would help them in the future was the goal of this teacher. Mrs. Flores had hoped to find more answers for getting students to care and work harder, and, in the process, get parents to be responsible and accountable along with the school system. She felt she was still searching for "magic" answers, but in the meantime, her good instructional practices in the classrooms worked, because she had the most success of all teachers in the district in bringing up test scores. Mrs. Flores said that at the level she taught, with at-risk children, she had more discipline problems than most other classes had. Even when her students were assigned to in-school-suspension (ISS), she still provided lessons and was accountable for those children. She said if schools could eliminate the discipline problems, and she felt they had gotten worse over the years, and if schools could figure out how to motivate apathetic students, teaching would be much easier. Mrs. Flores

believed that next to high expectations, discipline was the most important factor in the learning process (Interview 6, 11/7/03).

Table 13: Matrix of Findings for Discipline and Parent Involvement

Theme 6: Findings and Sources of Data Triangulation - Sources of Data: I = Interviews, S = Surveys, O = Observation, D = Documents	I	S	O	D
1. Discipline was a major key in the effective teaching/learning process.	x		x	x
2. Effective teachers make an intense effort to involve parents in school.	x		x	x
3. More and more parents have non-supportive attitudes with teachers.	x		x	
4. Teachers have varied ways of establishing good classroom discipline.	x		x	
5. Bad behavior caused constant interruptions and lost learning time.	x		x	

Notes: Document/Artifact for #1 = Appendix M
Document/Artifact for #2 = Appendix N

Theme 7: High Expectations for Students

The participants (P1, P2, P3, P4, P5, P6) of this study all set high expectations for their students, believing all students and subgroups could succeed and could perform well on achievement tests. These teachers knew how to engage the students effectively, remained committed to providing a classroom climate for learning, and worked to make students believe they could be successful. As these teachers deepened their level of good instructional practices,

they became quite skillful at raising the students' confidence in themselves. Most of the participants believed that high expectations, even more than discipline, were the most important factor in the teaching/learning process.

At whatever level she taught, Mrs. Blair (Interview 2, 10/6/03) shared that high expectations made a great deal of difference. If students were working below grade level, she made them believe they could make progress if they set goals and worked hard. Mrs. Blair expected a lot of herself as well as her students in order to get them working at grade level, using peer groups/teams to help lower-achieving children. She explained more thoroughly with those who did not understand or who had gaps in their skills. Mrs. Blair understood that at the end of the year, these students all took the same test, and even though she might have had several grade levels within her group of third graders, she stretched them as far as she could get them. She admitted to a certain level of frustration when she had the students working above their abilities because she did not want the student to be overwhelmed or feel ill-equipped. She balanced creating an atmosphere of high expectations with letting her students know she was there to help them. Mrs. Blair had the responsibility of several special education students in her classroom where students performed at different levels, yet these students were not tested on the level they performed. High expectations remained an important part of Mrs. Blair's teaching process with those students (Interview 2, 10/6/03).

With high standards for herself as well as her students, Mrs. Flores, (Interview 6, 11/7/03) believed high expectations to be the number one factor in the learning process. She expected her students to make an 80% or better on all skills assignments (Appendix J), or she re-taught the skill and the student did the assignment again. She said students did not like doing that, so they tried harder the first time. Students had to apply themselves in Mrs. Flores' classroom, and some had never been expected to do work that hard. Her goal for her students was to have everyone reading at or above the 80% level on the Oklahoma PASS test, and most of the time they succeeded. For at-risk and ESL students, that was a lofty goal, and she had students that she worked with one-on-one to convince them to reach that target. Mrs. Flores' apprehension was about the improvement level that the state raised each year until the year 2015 for NCLB. Her concerns were for her ESL students, at-risk students, and her own children. Her concern was "What is enough? When we get to the magic number, where to we go from there?" Mrs. Flores did not want the expectations from NCLB to destroy the expectation level she had worked very hard to build in her classroom (Interview 6, 11/7/03).

According to Ms. Edwards (Interview 5, 10/27/03), students would usually strive to meet a teacher's high expectations and even surprise themselves as well as the teachers. Ms. Edwards said expectations were immensely important, and that students, no matter what subgroups, should be performing on grade level with

"no excuses" in order to meet the mandates of NCLB. When Ms. Edwards expected more of her students, they had a tendency to expect more of themselves. She gave the example of a child who would start out by reading one book and then one more, soon realizing that they could read many books until it became a competition to see just how many books they could actually read. Often, Ms. Edwards said, they surprised themselves and ended up liking the competition (Interview 5, 10/27/03).

The three remaining participants agreed that expectations were very important and that they definitely had an impact on learning. Realizing that students are sometimes smarter than a teacher gives them credit for, Mr. DiCarlo (Interview 4, 10/27/03) said the more he expected of students, the more they understood what they themselves could really do. With expectations as important as discipline and maybe more so, Mr. Callahan (Interview 3, 10/15/03) said expectations were "everything" and if he had high expectations, students rarely disappointed him. If he told a student "Yes, you can do this," the student would invariably accomplish the task (Interview 3, 10/15/03). Echoing those statements, Mrs. Arbello (Interview 1, 10/3/03) said that if a teacher expected "nothing," he would get "nothing," but if a teacher expected "a lot," he would get "a lot." Mrs. Arbello said teachers got exactly what they expected; however, students were prone to forget, quite unintentionally, and that teachers should not expect them to

be adults and should expect to remind children of what they had forgotten (Interview 1, 10/3/03).

Table 14: Matrix of Findings for High Expectations of Students

Theme 7: Findings and Sources of Data Triangulation Sources of Data: I = Interviews, S = Surveys, O = Observation, D = Documents	I	S	O	D
1. Effective teachers had high expectations of students.	x		x	x
2. Students performed better if they believed they could be successful.	x		x	x
3. Expectations were the number one factor in the teaching/learning process for these participating teachers.	x		x	

Note: Document/Artifact for #1 & 2 = Appendices J & O

Other Issues or Non-Issues that Effect Student Achievement

The researcher asked every participant if they had the necessary resources and district support they needed to meet Adequate Yearly Progress. Every teacher answered in the affirmative. As a group, these teachers (P1, P2, P3, P4, P5, P6) thought having enough time and unrealistic expectations were the major issues in effecting student achievement. The researcher anticipated the issue of money or the Oklahoma state teachers' salary schedule to arise during all the interviews, but that did not occur. Not one teacher mentioned the word "money" or asked for any additional resources to comply with the mandates. They felt the problems rested elsewhere.

Whether it was from the NCLB Act or state mandates, these six teachers (P1, P2, P3, P4, P5, P6) understood that change was a necessary process for any time of school improvement. They remained very positive in their thinking and wanted to be a part of the solution rather than the problem. Their positive attitudes affected the way they went about the teaching/learning process, as compared to the negative attitudes that always surfaced in a change process. These teachers had no excuses for their efforts, and very few excuses for their students; they simply went about finding ways to meet the mandates. One teacher (Interview 6, 11/7/03) worried about the deterioration of students' motivation, but she wanted ways or answers in raising the students' motivation. Other teachers were concerned about student behavior; however, they felt that establishing classroom discipline was a part of their jobs.

These teachers (P1, P2, P3, P4, P5, P6) felt that the district leadership had encouraged and supported their efforts in preparing for NCLB and AYP. Mrs. Flores (Interview 6, 11/7/03) said the district was well-prepared for NCLB with a total district alignment and district assessments that were aligned to the PASS standards. Teachers received some professional development training from the curriculum director in using data. Each building site also had technology, including software, in place that would assist in preparing students. All six teachers (P1, P2, P3, P4, P5, P6) felt that this district was ahead of others in meeting the mandates, and they had begun to meet in grade level meetings at their

sites to discuss all data they could collect. Encouragement from their principals or other district administrators occurred on a frequent basis in this district..

In addition, these teachers thought that some of the Oklahoma Academic Performance (API) requirements were unfair. According to Ms. Edwards (Interview 5, 10/27/03) one elementary school that served a military installation had a very high mobility rate, with students moving in and out of the district frequently, and the site did not get to count the students unless they had been present at that site for a full academic year. The counselors and teachers had to go back to last year's register to see who and who had need been at that site for a full year, meaning one full year from this year's test date (Interview 5, 10/27/03). This not only meant more paperwork for teachers, they believed that it was unfair that they did not get to count half or more of their students. Teachers from other sites in the district also thought they should be able to count the students that were well-prepared for the test, and that Oklahoma's rules were unfair. These teachers wanted to count all students who had been in attendance for the current physical year, but they were not allowed to do that.

One issue brought heated responses from the participants, and that topic was about teacher involvement in the federal and state mandates. These teachers (P1, P2, P3, P4, P5, P6) said they were not consulted in the design or implementation process of NCLB or the state API, and they did not know any teacher who was. They knew NCLB was patterned after the Texas TASS system,

but they did not know how much teacher involvement in that process had occurred. These six teachers (P1, P2, P3, P4, P5, P6) did say that if effective teachers had been intricately involved, the design and implementation of NCLB would have been much different. They felt there was no discussion about parental responsibility or involvement, highly mobile and multi-marriage families, programs for one-on-one tutoring of students, extra help for the economically disadvantaged, the variance in special education rules and regulations, and realistic expectations. These teachers (P1, P2, P3, P4, P5, P6) believed they were the ones who worked every day in the "trenches" and that they should have been consulted.

CONCLUSIONS AND RECOMMENDATIONS

Chapter V

Introduction

In researching the effect of individual teachers, attempts to measure their effectiveness in a quantitative way have brought forth varied results (Marzano, 2003). However, according to Marzano, most researchers would agree that the impact of decisions made by effective teachers was "far greater than the impact of decisions made at the school level" (p. 71). Studying the impact of highly effective teachers in one Oklahoma district and the methods they have implemented to be successful could make an impact in other classrooms. What Marzano (2003) said could be measured quantitatively was that on the average, effective teachers' students showed gains on the average of 53 percentage points in one year, while a less effective teacher's students showed gains of around 14 percentage points in one year.

Data gathered from "highly effective" teachers' practices could dramatically improve the teaching/learning process for students, an ongoing endeavor for almost all teachers. In addition, effective teachers will usually share what works and what does not, along with what they consider "realistic" and "unrealistic" expectations from students. The educational community should pay heed to these "highly effective" teachers' conversations.

The conversations from the six teachers in this Oklahoma district with this researcher reflected very positive thinking about their students and their classrooms but very guarded opinions about the No Child Left Behind Act (NCLB), Adequate Yearly Progress (AYP), and the Oklahoma Academic Performance Index (API). Those guarded opinions were not totally negative, because teachers wanted every child to succeed, just as the NCLB Act mandated. However, the six teachers interviewed all (P1, P2, P3, P4, P5, P6) felt that the expectations from NCLB and the Oklahoma API were not only unreasonable but also inequitable in many instances, with the state and federal governments failing to provide the necessary tools for improvement.

This case study, both qualitative and quantitative, endeavored to answer the following research question: How had "highly effective" teachers found methods to use Oklahoma achievement test results and other educational data to impact classroom instruction and improve learning that would ultimately result in "adequate yearly progress" for No Child Left Behind.

The purpose of this case study was to observe experts in their field, "highly effective" teachers who will lead other educators "to find their way into the current reform debate" (Meier, 1995, p 184). Highly effective teachers learned from each other and they learned by trying new things for themselves, and those new to the field could also learn from those wise experts. These were the teachers who worked diligently to provide students with tasks and activities that

were inherently engaging (Marzano, 2003, p 149), and then provided students with the feedback they needed to be successful.

This section will report the conclusions and recommendations from six highly effective teachers in one Oklahoma school district (P1, P2, P3, P4, P5, P6) and the data collected from the case studies. As the data were organized into themes from the transcribed interviews, the analysis of and recommendations from that information followed the same pattern.

The Initial Survey

From the initial survey, 30 teachers indicated their district had a well-defined curriculum, had adequate district assessments, and had sometimes looked at state test blueprints and achievement results. Those 30 teachers used their district and classroom assessments to improve instruction and judge student progress. They planned instruction according to the PASS standards and the state test blueprint. However, the teachers had not used the state achievement test results in a meaningful manner, relying instead on district pre/post tests and their own classroom data.

These 30 teachers also felt that state and federal governments had not provided adequate resources, namely valid, reliable data, to meet the mandates of NCLB and AYP. Teachers did not believe that the federal and state governments had involved them in the design of NCLB and the Oklahoma API, nor had teachers anywhere been involved in the design. They also did not feel that

teachers had a part of designing the state achievement assessment system, else teachers would have a greater say in who was tested and at what level. In addition, the state test results or reports that teachers received were not timely nor were they easy to read and understand. Those sentiments were echoed by the six key participants that were chosen to be interviewed.

The themes that emerged from the interviews of six key teachers in the third grade (NRT data) and fifth (CRT data) grade included (1) issues with the Oklahoma State PASS Standards; (2) questions about Oklahoma assessments and blueprints; (3) the sharing of teachers evaluation and use of data; (4) teachers' proven practices and classroom interventions; (5) the lack of time for instructional planning; (6) issues of discipline and parental involvement; and (7) teachers' high expectations for students. Additional topics, or "Other Issues or Non-Issues that Effect Student Achievement," surfaced from the interviews, relative issues minimally addressed by the respondents.

Theme 1: Conclusions about Oklahoma State PASS Standards

All of the six "highly effective" teachers (P1, P2, P3, P4, P5, P6) chosen for interviews used the state PASS standards to plan instruction; however, they felt the state had outlined far too many standards and were unrealistic in their expectations. The teachers' daily lesson plans reflected their practice in documentation of the standards. However, they felt their district was ahead of

other districts because they had worked on a total K-12 curriculum alignment in all subjects and had spent three years on this process.

The alignment process these teachers worked through had made them painfully aware of how broad the PASS standards were, nor only in reading and math, but in all subjects. All participants were pleased they had aligned the curriculum and had become so familiar with the standards; but they found many skills in all subjects that they were expected to teach, including reading, math, science, language arts, social studies, music, art, health, physical education, and some introduction to a foreign language. The PASS document for elementary teachers was indeed a formidable document, causing these teachers to feel inadequate and stressed.

The teachers felt stretched just to teach basic standards in depth. If a teacher realized that students had not mastered a concept and had to stop and re-teach, their timeline for all the other standards was limited. Teachers documented the teaching of their PASS standards in various ways, with their lesson plans, hand-made lists, or district curriculum guides, but they believed that there was not enough time to adequately teach all of the standards for the entire core curriculum. Some felt that since reading and math were the two subjects included in the Oklahoma API, that they felt pressured to spend more time on those skills. The Oklahoma State PASS standards should be streamlined with only the most basic skills and concepts included for all core subjects.

Test blueprints were only somewhat helpful to teachers; however, the six teachers interviewed would all request that the tested standards to be well-defined with information to indicate the exact skills and objectives taught in the various categories. They called the categories vague with little information attached to the blueprint. Mrs. Flores (Interview 6, 11/7/03) said that teaching the category "literary elements" was like a "stab in the dark" because there were many different lists of literary terms, and the test might only include a few items.

The report that came with the norm-referenced tests for the third grade included a list of skills (a blueprint); the teachers (P1, P2, P3) also felt that list was vague, asking exactly what was included in "problem solving" or "listening skills" or "textual reading." If the state would provide more information on the categories for the teachers, they could better prepare their students, not feeling so stressed to cover a massive amount of material before test time. To do their jobs adequately, the state should consider providing more information on the blueprint categories for the teachers, rather than have the information remain a "deep, dark secret" as Mrs. Flores (Interview 6, 11/7/03) charged.

Recommendations, Theme #1: State policymakers might consider

1. Reviewing PASS standards to be streamlined in all core curricular areas with only the most basic skills and concepts included and selecting fewer skills in each subject to be tested, especially in the fifth grade where there

were standards in reading, math, science, social studies, and in writing (language arts).

2. Providing more information on the blueprint categories and subgroups, such as problem solving and giving teachers some indication, as precise as possible, of the skills to be tested.

Theme 2: Conclusions about Oklahoma State Assessments and Blueprints

The teachers interviewed (P1, P2, P3, P4, P5, P6) were tired of the changes in the Oklahoma State Assessment system. As one participant said, "tests change, test companies change, and guidelines change" (Interview 1, 10/3/03). With the system in place in 2003-04, elementary teachers administered a norm-referenced test (NRT) at the third grade and a criterion-referenced test (CRT) at the fifth grade.

In the last decade, the state had changed tests and test companies numerous times (see Table 1). At the beginning of the 2003-04 school year, Oklahoma teachers were given a chart of further changes that could be anticipated within the next five years. One of those changes was to drop the NRT at the third grade level and replace it with a CRT; another change was to test at least the subjects of reading and math from the third through the eighth grades as mandated by NCLB (OSDE, Student Assessment, 2003). With the assessments that were in place, teachers felt frustrated keeping up with the constant changes; and with further changes in the future, they may continue to express dissatisfaction.

The six teachers interviewed had few comments on sanctions that might occur if they failed to make AYP; however, they disagreed with the state testing all children at the same grade level with the same test. Their concerns were for the special education and ESL students; each of those subgroups was included in a safe harbor by the state, but teachers said that the safe harbor was not nearly enough. All teachers believed that this year's special education or ESL students being compared to last year's students would not be productive. These teachers thought this practice to be implausible and went against everything they had been taught, namely that these students' growth or gains should be compared only to their own charts and records. Moreover, one year's students might be average or almost average, and the next year's students might be far below that.

With the safe harbor clause, a building or district did not meet AYP if this year's students were 10% below last year's students (OSDE, 2003). Teachers referred to the Individual Education Plans (IEPs) and Reading Sufficiency Plans (RSPs) they had in place for students. With an IEP or an RSP, a teacher tested the student's performance level and adapted learning materials to fit the learning level of the student. If teachers were asked to prepare all their students for the third and fifth grade tests, on grade level, including students on an IEP or RSP, they would be asked to do the impossible. These special needs students had already been identified as in need of extra help or below grade level, or the teachers would not have a plan for them.

The state of Oklahoma was in the process of designing an alternative assessment for special education students, but that assessment could only be used for 1% of the district's students. These teachers (P1, P2, P3, P4, P5, P6) did not feel the safe harbor clause or the alternative assessment would help in the overall picture of meeting adequate yearly progress. What they wanted most was for special needs students to be tested at their performance levels.

To be most effective, these teachers (P1, P2, P3, P4, P5, P6) asked for current data on the students they had in the current year. District group reports with test data could be handed to the next level of teachers, except that the mobility rate for all schools was higher than it had ever been, and teachers would not have data on many students. Also, the test results would not indicate the students' retention of skills over the length of a summer. Instead, teachers relied on the district's pre test or their own assessment or placement instruments at the beginning of school. They felt these were far more accurate than the state tests where students could have had test anxiety or lucky guesses.

These teachers also shared that using last year's data to improve this year's students was like what one participant called "comparing apples to oranges" (Interview 2, 10/6/03). Participants wanted a way to test students at the beginning of the year, showing their progress throughout the year, and they wanted reliable data on this year's current students.

Methods of reporting the state data made the results difficult to read and understand for these teachers. In the last two years, 2002 and 2003, the state included multiple reports for all subgroups in the site level reports, and all participants did not know which pages contained the information they needed. They also did not comprehend how the information or data from those reports were then translated to the Oklahoma API formulas. Then, there were differences in the way norm-referenced tests results and criterion-referenced tests were reported (see Table 3, Chapter 1, page 48) complicating matters for teachers who wanted to interpret the scores or results that were plugged into the API formulas. If teachers were expected to use state data to meet AYP for the Oklahoma API, they needed reliable, current data they could understand, data they could use to estimate gains or losses for progress on the API.

At the time of the interviews, these six teachers (P1, P2, P3, P4, P5, P6) used very little of the state data to improve instruction. More importantly, they used their own assessments to track student progress and to plan for the mandates of AYP, consulting the classroom data more than they did any state data, primarily because that was data they could understand. Using their own assessments made it possible for teachers to have current year data on their current students; however, it created a hardship for many teachers to have to create their own pre tests, skills tests, and other diagnostic tools. Even practice tests available on the state website were inadequate for this group of teachers,

with them still having to make their own practice tests. With the classroom assessments, both formal and informal, those that teachers had already built into their system, they believed that testing and assessment should be a part of everyday instruction, not a once-a-year item.

Recommendations, Theme #2: State policymakers might consider

1. Designing an assessment system that would eliminate the changes in tests and test companies every year or two, allowing teachers longitudinal data from the same test vendor and test banks to identify trends.
2. Testing ALL special needs students at their identified performance level, and not have just 1% of the students tested with an alternative assessment.
3. Giving teachers reliable data on this year's current students and let them track the individual progress a student has made from one year to another.
4. Providing teachers information and data for AYP and the Oklahoma API in data and terms they could understand.
5. Including assessments other than the once-a-year test, especially for some students and using a combination of district data or portfolio results collected and reported to the state by the district.

Theme 3: Conclusions on Teachers' Evaluation and Use of Educational Data

Teachers in this district (P1, P2, P3, P4, P5, P6) wanted to become data-driven and base their educational decisions on more than checklists or instinct; however, using the state data was difficult for them. Some professional

development had been provided the teachers, and they thought that had been helpful, but the training still had not convinced them that the state data were valid. Their first statements to the researcher about the state assessments included comments about bias and validity.

One participant (Interview 3, 10/15/03) remembered a question that a student had asked him about after the test. The test question had a picture with a house in front and an oil well in the back yard, and the question asked what, if anything, was wrong with the picture. This participant's Oklahoma third grade students had said "no," because they saw oil wells everywhere, including in back yards, as quite normal. This participant (Interview 3, 10/15/03) said he knew that most of his students got the answer wrong. Going further, the teacher wanted to know why the students were being asked such a question because he thought that it had nothing to do with an actual core "skill" that students would use in the future. Educators should be assured that core skills measured on the state achievement test were valid and unbiased.

All of the participants (P1, P2, P3, P4, P5, P6) agreed that their own data from district pre tests and classroom assessments were more helpful than state data provided at the end of the year. Diagnostic and placement tests were the norm for these six teachers, with much of that data transformed into charts and graphs; and as routine, assessments were an integral part of these teachers' classroom practices.

These teachers used the state test results only for a guideline, relying on oral reviews, demonstrations, hands-on activities, and paper/pencil assessments to give them valid, reliable data. Mrs. Blair (Interview 2, 10/6/03) said that in one year, the state tests might show that she did an "awful" job of teaching, but by her own data, the students might have made tremendous progress throughout the year. Mrs. Blair (Interview 2, 10/6/03) wanted the state to test her students at the beginning of the year and then again at the end of the year. All six teachers (P1, P2, P3, P4, P5, P6) agreed they should be responsible and held accountable for their year of teaching, but all argued that using state test results from the end of one year and comparing them to the next year was not the way to achieve the task.

Teachers looked at the state achievement test data at the end of the year, or at the beginning of the next year to see what areas were "low" or needed improvement; however, these six teachers stressed several times that this was a once or twice a year practice with them. All the charts and graphs that came with the achievement tests were not used because teachers firmly believed that the state data were only cursory information to their task at hand, which was the next set of students.

The teachers admitted that they were always interested or curious in the test results and what general areas might have been low, but they still attached very little importance to the state test data. These teachers were asked if that practice was from years of habit or if they truly believed that the state data were

of little assistance. Truly, they said they did not believe the state test results were a reflection of their program of instruction, because the test question content was from a vast set of standards and such broad areas and that predicting what would be on the test was an impossible task. These teachers wanted state achievement tests that measured a streamlined set of standards, a set of core skills that all students should know and be able to accomplish.

The participants in this study (P1, P2, P3, P4, P5, P6) wanted practice tests that would replicate the state tests, giving teachers and students some information on how to prepare for them. The teachers had visited the state website that provided practice test items, but they said the information took too long to download on their computers, and when they could get the information, it was not useful. The test questions were for easily recognized skills in easy to answer formats. These participants had made their own test bank of questions, using the district pre test item analysis, identifying low skills, making up test questions for practice on the students' substandard areas. They said that even the practice tests they ordered from test companies covered a minimal set of skills, rarely matching what would be included on the Oklahoma state tests. If the state of Oklahoma could provide a release test, as Texas did, teachers would not have to waste so much time making their own practice assessments nor would they have to guess which skills and objectives would be tested.

Recommendations, Theme #3: State policymakers might consider

1. Assuring educators and students that test items from state achievement tests were valid and unbiased, that test items measured core skills, and that there were limited skills measured for any one grade level.
2. Measuring only well-defined basic skills and objectives from the standards of core subjects, so teachers did not have to guess what would be tested.
3. Providing teachers with a practice test or release items on current skills and objectives in accurate format.
4. Using state test results from the end of one year, comparing them to the next year's set of students, might not indicate AYP for that set of students.

Theme 4: Conclusions on Proven Practices and Classroom Interventions

The six highly effective teachers (P1, P2, P3, P4, P5, P6) modeled practices and interventions in their classrooms that made a difference in children's lives and created various learning opportunities. Those proven practices had the greatest impact on student learning, outweighing the use of state data in improving instruction. Each of the participants had his/her own strengths in the teaching/learning process, but several practices emerged as common to all. Those practices should be emulated by teachers everywhere, as they worked for this set of six teachers. However, the practices took large chunks of teachers' time and they took total dedication on the part of the teachers. Those proven practices are outlined in the Table 15 below.

Table 15: Proven Practices of Highly Effective Teachers

1. Item Analyses	Teachers write their own item analyses from district or teacher-made diagnostic, placement, or pre tests.
2. Teacher-made Assessments	Teachers tested students in various ways, including diagnostic, placement, or skills test, both formally and informally
3. One-on-One Tutoring	Time limited help for struggling students, and teachers stayed after school on their own time, helping students master concepts.
4. Cooperative or Group Work	Group work gave stronger students the opportunity to help other students, and oral reviews helped new students.
5. Demonstrations and "hands-on"	"Hands-on" activities were very helpful to most students in the learning process; however, those activities take up more valuable time.
6. Guided Practice	Teachers watched students closely to give immediate feedback and interventions.
7. "Looping"	Cumulative reviews reinforced skills so students could transfer learning into long-term memory.
8. Graphs, Charts, Timelines	If students can visualize or track their own progress, the student has immediate feedback in a meaningful way.
9. Practice Tests	Students were taught test-taking skills along with the skills practice, review, and reinforcement from the exercises.
10. Technology	Students benefit from diagnostic and remedial software, and they like using the computers.
11. Portfolios	Portfolios or folders give immediate documentation on student work, for the student, parent, and the teacher.
12. Use of Other Resources	Teachers frequently supplemented textbooks with other materials, including lower and higher levels for remediation, and enrichment.

Recommendations, Theme #4: All teachers might want to consider

1. Creating their own classroom data, including diagnostic, placement, and pre tests. Teachers need the know-how and resources, such as computer software, to accomplish this. Those who indicated they did not know how

to build diagnostic assessments should be offered training to create data and databanks.

2. Asking state policymakers to consider funding for one-on-one tutoring programs before or after school. These were the resources missing from the federal mandates in NCLB, along with the funds to hire extra teachers for pullout tutoring programs during the day.
3. Incorporating the use of demonstrations, "hands-on" activities, cooperative work, and guided practice into their everyday classroom activities, often with supplemental resources.
4. Learning to "loop" (review and test cumulatively) skills throughout the year for maximum retention.
5. Using graphs, charts, and timelines to provide students with valuable instructional and learning tools.
6. Implementing practice tests and technology such as computer software to provide needed reinforcement of skills and diagnostic information.
7. Keeping portfolios or folders of important work on every child to provide valid data and documentation.

Theme 5: Conclusions on Issue of Time for Planning Instruction and Using Data

"Time" proved to be the significant element that enabled these six "highly" effective" teachers (P1, P2, P3, P4, P5, P6) to be as successful as they were. This group dedicated chunks of time outside the school day in order to

create and administer the proven practices (see theme 4 above) that helped their students learn what they were supposed to learn. Working long hours during the school year and in the summers was a common practice that these teachers took for granted, and they continually exhibited positive attitudes about those long hours, the extra time a part of the high expectations that had of themselves.

The six teachers found the time to create new, innovative ways to teach and assess their students, making their classes interesting. The district assessments and placement tests they administered at the beginning of the year took away instruction time, but these teachers found the item analyses that came with the district tests to be quite valuable. Spending time after school, the teachers took the item analyses and made lists of skills, from "weak" to "strong," using those for timelines in which to teach the important standards.

All of the teachers (P1, P2, P3, P4, P5, P6) conceded that "hands-on" activities, demonstrations, and oral reviews took time; however, students seemed to retain the skills much better so these teachers were reluctant to simply assign lessons from the text as many teachers did. One of the participants (Interview 2, Field Notes, 10/6/03) made the statement that anyone could walk in off the street and assign the exercises from the textbook; however, she took the word "teach" seriously, making sure that she had taught the assignments, even vocabulary words that were easy to "assign." All of the other proven practices listed in Theme 4, Chapter 4, took time away from grading papers, planning lessons, and

other school paperwork, but these teachers (P1, P2, P3, P4, P5, P6) were willing to take the necessary steps to make sure their students were prepared properly.

If these teachers (P1, P2, P3, P4, P5, P6) could change the way school operated, they would design a schedule with maximum instruction time and limited interruptions. They understood the necessity of assigning students for extra time in special programs, realizing that music or art or sports may be a vital outlet for unmotivated students; still, these classroom teachers wanted long periods of time, an hour and a half at least, for reading and for math, every single day. These teachers did not ask for extra "time" during the school day for lesson planning for themselves; they asked for the necessary time for actual instruction with no interruptions.

Seeing "time" as a precious commodity, these teachers wanted their principals and other teachers to help them design schedules that would allow for the "hands-on" teaching and guided practice. These teachers asked for understanding from parents, not wanting students pulled out of class unless absolutely necessary, and they asked for understanding from their principals, wanting assemblies and pep rallies on a very limited basis. They insisted that instructional time should be guarded at all costs, even eliminating the intercom announcements during the day. They wanted the time to adequately teach the state standards for the benefit of the student.

Finding that time did not allow them to get all of the standards taught by test time, the participants (P1, P2, P3, P4, P5, P6) identified the core skills that could be mastered. They used their item analyses and other diagnostic tests to detect weak skills, starting on those at the beginning of the year. Teachers used the state test blueprints to see how many questions were assigned to which group, using that information to decide how much time to spend on that set of skills. These six teachers all (P1, P2, P3, P4, P5, P6) said that teaching the standards and preparing for the test left little time for anything else, feeling pressured and stressed until after the state test.

One participant (Interview 5, 10/27/03) said that teachers had no choice but to teach test skills and test preparation if they wanted their students to be successful on the state exams. Another participant (Interview 4, 10/27/03) said that teachers did not relax and have any "fun" with their students until after test time. These teachers were at the point of being "overwhelmed" with two of the participants (P3, P4) using that exact word. Reducing the number of state standards and allowing teachers to teach the skills in depth should be better for students and teachers alike, lessening the stress and pressure throughout the year.

Recommendations, Theme #5: Successful teachers might consider

1. Dedicating chunks of time outside the school day in order to implement the proven practices of these "highly effective" teachers.

2. Taking the word "teach" seriously, refusing to simply assign lessons from the text books or work books.
3. Asking principals and other school leaders to guard teachers' instructional time, making good schedules a priority.
4. Requesting the OSDE reduce the number of skills under the standards in each core area, allowing teachers the time to teach core skills in depth.

Theme 6: Conclusions on Discipline and Parent Involvement

All teachers interviewed (P1, P2, P3, P4, P5, P6) believed that discipline was a major element in establishing an effective teaching/learning environment. They worked to establish good communications with their students' parents, realizing that parental involvement was the best way to help establish good discipline. Each of the participants had a strength in gaining the students' respect, and one teacher's strength did not outweigh another's in the process. What emerged from their conversations was that any teacher must find his/her process of establishing discipline before the important learning process could begin.

These "highly effective" teachers (P1, P2, P3, P4, P5, P6) believed that good behavior in schools had eroded in the past years, taking upon themselves the task of teaching self-control and self-discipline. Although these teachers would like for parents and guardians to have their children ready for school, they worked diligently to gain proper control of their classrooms. They believed very strongly that they could not teach effectively with disruptions. One disruptive student

could take several valuable minutes away from other children's instruction time, and if several students misbehaved, little instruction would be gained. These teachers labored to make their classes interesting while students learned the necessary skills, and one participant (Interview 3, 10/15/03) said that he would retire soon if the discipline problems did not become better. All of these teachers asked that parents become more actively involved in their children's school day, believing that teachers had too many tasks already.

Recommendations, Theme #6: Teachers might consider

1. Insisting on good behavior from all students, thus allowing the teacher time to teach students who behaved well and wished to learn.
2. Working with parents to take responsibility for their children's behavior and to assist the teacher in establishing self-discipline from their children.

Theme 7: Conclusions on High Expectations for Students

The participants of this study (P1, P2, P3, P4, P5, P6) believed that high expectations of students were just as important as good classroom discipline, and perhaps the number one element in the teaching/learning process. Even though these teachers believed that all students could succeed, and worked to make that happen, they knew that students should be working on their own ability levels. All six teachers worried that testing all students at the same grade level with the same test was against everything they had ever learned and would eventually

destroy the confidence they had built in their students. This dichotomy seemed to be the conflict that led to the most frustration from these teachers.

Each of these teachers had learned or knew instinctively how to engage the students effectively, becoming skillful in raising a student's confidence. Expecting a great deal of all their students, they (P1, P2, P3, P4, P5, P6) made students believe they could progress if they set goals and worked hard. Two of the participants (P5, P6) asked their students to make an 80% or better on all their work, whether that was homework or exams, and if they did not, they were asked to repeat the process with the teacher's help. Usually wanting to please their teachers and discovering that they could make good grades, these students strived even harder to complete the work correctly the first time. The students' improved confidence also carried over into other subjects for other teachers. However, these teachers knew the ability level of their students and knew which work they could complete successfully and which work they could not.

These teachers (P1, P2, P3, P4, P5, P6) "stretched" their students as far as they could, and they did not know of any teacher who wanted to leave a child "behind." They worked even harder with subgroups that were not performing well on the state tests; however, they believed those subgroups were already being pushed too far. These teachers said there were various reasons some groups of children were not performing well, including language barriers, poor home lives with no support, learning disabilities, and absence of reading materials at home.

With high expectations, these students could progress through the year, but still not be able to reach the state standards.

Recommendations, Theme #7: Teachers might consider

1. Working to raise more students' levels of confidence with high expectations.
2. Maintaining high expectations of all students, but those expectations should be at a realistic level.
3. Continuing to work with and test students on their individual levels of performance.

Topic 8: Conclusions on Other Issues/Non-Issues on Student Achievement

Having anticipated that the six teachers interviewed (P1, P2, P3, P4, P5, P6) would at some time bring up the subject of money or Oklahoma teaching salaries, this researcher soon learned that these teachers were interested in sharing much more important issues. Neither of the words "money" nor "salaries" was mentioned at all in these conversations. However, this district had paid teachers above the minimum state salary schedule for several years, with funding from military impact aid helping with the district finances. These teachers would have liked for after-school programs to be offered for one-on-one tutoring, but they did not discuss the funding for those programs. During the interviews, these teachers were also asked if they had the necessary resources to work toward adequate yearly progress, and they all answered in the affirmative. These answers from

these six teachers should send a strong, convincing statement for those interested in educational funding – these six effective teachers were not asking for more money, they were asking for more time and realistic expectations. Granted, additional funding might be needed to gain the instructional time that teachers needed, but realistic expectations would have to be modified from the NCLB Act and most certainly the requirements from the Oklahoma API.

These teachers (P1, P2, P3, P4, P5, P6) asked that the state of Oklahoma take a second look at the rules and regulations surrounding the API. First, representatives of the assessment department of the Oklahoma State Department of Education (OSDE) had said that the rules surrounding a full academic year were meant to help schools, not harm them (OSDE, 2003). However, a high mobility rate for many schools in this district would decrease test scores, especially for those students on a military installation. At that elementary school, more than half the children in any grade level might be new in August when school started. When teachers asked OSDE test representatives if they could count those students anyway, they were told "no." (OSDE, 2003) The second element that state legislators should consider exploring further is the safe harbor clause that compared one year's special needs student to the progress of the previous year's special needs students. With so many ESL students with language barriers and so many different learning disabilities among our students, those test scores will surely fluctuate because those students were being given the

same test as all the other students. These were the students who would more easily "freeze" at test time or make "lucky" guesses, and if we want valid, reliable test data on these students, the present system would not succeed.

Teachers who expressed these statements were not making "excuses" for not helping the lower-level students; these were statements from our most highly effective teachers (P1, P2, P3, P4, P5, P6) that worked with these students every day. These were the teachers that policy makers, especially those who make the rules for the Oklahoma API, should consider their suggestions.

Recommendations, Topic #8: State policymakers might consider

1. Hearing from effective teachers who had asked for more "time" and "realistic" expectations, not for more money or special funding.
2. Taking a second look at the Oklahoma Academic Performance Index, especially regulations concerning "high mobility" and "safe harbor."

Implications from Other State Systems for Oklahoma School Leaders

Oklahoma had spent hundreds of thousands of dollars over the last few years putting the PASS standards and criterion-referenced assessments into place. Oklahoma had also struggled with how to meet the accountability of AYP and putting into place the Academic Performance Index. There were taxpayers who asked why they had to spend enormous amounts of money to replace what they did not consider "broken" (Popham, 2002) when almost every state in the nation faced budget crises over the past few years. They considered it an unnecessary

waste of money to emulate what one state (Texas) had in place and was evidently not working well.

When meeting the mandates of NCLB, Oklahoma had to decide whether to devise separate assessment measures or blend the new with the old. Nebraska appeared to have the edge over other states because it had not tried to layer a new accountability system over an old one (Chapter 3). Since Texas already had its system in place, there were few major adjustments, but they had implemented a new TAAS II that measured their Texas Essential Knowledge and Skills (TEKS). Oklahoma had contended with making minor adjustments in an Academic Performance Index and did not completely overhaul the whole system, blending new assessments with old ones.

States, including Oklahoma, could have used locally developed criterion-referenced assessments, and they did not have to absolutely follow the Texas Model. NCLB guidelines for standards and assessment allowed states to use either criterion-referenced tests or "augmented" norm-referenced tests. Many of the studies published in the last two years encouraged states to consider placing more of the responsibility for assessment at the local level, much like that accomplished in Nebraska and Vermont, and less at the state level where the predictable multiple-choice achievement tests were used (Amrein, 2002: Berger, 2000; Davies, 2000; Meisels, 2003). With only ten states reporting the use of supplementary criterion-referenced tests, one study (Meisels, 2003) suggested that

states consider doing just that because the principal purpose of those types of tests were to identify students in need of intervention and to determine appropriate instructional strategies. Meisels reported a study where the Iowa Test of Basic Skills (ITBS) was used to measure the progress of low-income, urban, elementary children who were enrolled in classrooms where a Work Sampling System (WSS) was embedded. The WSS was a curriculum-embedded performance assessment used for at least three years. Results for the WSS group reported growth in reading that exceeded the demographically matched contrast group. Children in that group also made greater gains in math. The study suggested that states should consider standardized tests at the state level to be used in conjunction with a curriculum-embedded assessment that would enhance teaching and improve learning. At the same time, these results showed that the curriculum-embedded assessment should also increase state test scores.

With a system of accountability that included performance-based assessments embedded in the curriculum, Stiggins (2001B) reported that test scores would also rise if the students were involved in the day-to-day assessment process. When students were involved in a non-threatening assessment process, they would also take more responsibility for what they were required to learn and what they had not learned. Both standardized achievement tests and local diagnostic type tests could affect learning, but the process would need to be well-constructed and manageable by teachers. Most classroom teachers have heard the

question "Will this be on the test?" Adults in college classes could and did ask that same question. If students know the material will be on the test and that they have to learn a concept, most will attempt to take control of their own learning.

Final Recommendations

From information gathered on what had worked and what had not worked from the survey of 30 teachers and in-depth interviews with six "highly effective" teachers, Oklahoma policy makers, along with educators, might consider the following statements:

- One assessment or achievement test, according to NCLB, did not have to be used as the only instrument in ranking how well a school or district performed from one year to the next.
- A combination of state and local assessments, including performance-based and portfolios, could be used in determining the proficiency of schools and districts on the Oklahoma Academic Performance Index.
- "Highly effective" teachers embedded curriculum assessments in everyday instruction, to measure progress and guide their next steps. Teachers and students were very involved in that assessment process.
- The general public could make simplistic interpretation of test results as they were combined with subgroups and demographics, not realizing there was no "causal" evidence for any subgroup that could be gathered from one test.

Recommendations for Further Study

Public education could offer students in the twenty-first century both quality and equity, especially with technology bringing all sorts of knowledge to the classroom. Technology had made it easier for the teacher and the student to have an abundance of materials and resources used in the teaching/learning process. With a wealth of information at their fingertips, graduating students should be more knowledgeable and more prepared for the world than ever before. Public education, according to Marzano (2003), stands at a crossroads where it could "enter the best of times" rather than the "worst of times." Following the advice of wise experts of "highly effective" teachers would be "exceedingly wise" and ignoring it would be "exceedingly foolish" (p. 178).

However, educational excellence could not be gained with standards and standardization, or as McNeil (2000) called the "logical consequences of the system that "de-skills" teachers and education (p. 270). McNeil believed that schools should encourage teachers to bring their own best knowledge to the classroom and create environments where all children could bring their strengths and capabilities to learn. Standards should indicate that all students ought to be learning the same thing, not that they should all be just alike, or as one participant (Interview 4, 10/27/03) in this study called from the same "cookie cutter" and able to pass the same test at the same grade level.

Because the "stakes" became so high for Oklahoma school districts with an emphasis on reading and math test scores in the Academic Performance Index, additional research should be performed to identify the negative consequences of ranking schools. With that continuum in place, there would always be one district at the top and another one at the bottom, but would that be the most effective way to measure Oklahoma schools? Would the "high-stakes" testing continue to take the "fun" and "creativity" out of school for both the teacher and the student? Would the tests continue to affect course content and teaching practices in adverse ways that have not been identified? Would district rankings of Oklahoma schools diminish the highly individualized practices of "highly effective" teachers (P1, P2, P3, P4, P5, P6), ones like those interviewed for this body of research?

REFERENCES

Academic Performance Index (2002). Oklahoma State Department of Education. Available: http://www.sde.state.ok.us

Achieve (2002). *Measuring up: A standards and assessment benchmarking report of Oklahoma.* Achieve, Inc., August, Washington D.C.

Amrein, A. & Berliner, D. (2002). High-stakes testing, uncertainty, and student learning. *Education Policy Analysis Archives.* Available: (http://epaa.asu.edu/epaa/v10n18.html).

A Nation At Risk (1983). National Commission on Excellence in Education. Available: http://www.ed.gov/pubs/NatAtRisk/risk.html

Anfara,V., Brown, K. & Mangione, T. (2002). Qualitative analysis on stage: Making the research process more public. *Educational Researcher*, 28-38

Bainbridge, W. (2003). Texas model for school achievement doesn't hold up. *The Columbus Dispatch,* April 16, 2003.

Balanced Assessments (2004). Harvard Graduate School of Education. Available: http://balancedassessment.gse.harvard.edu

Berger, J. (2000). Does top-down, standards-based reform work? A review of the status of statewide standards-based reform. *NASSP Bulletin*, 84(612).

Black, P., & William D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, 80(2), 139-148.

Bodgan, R. & Biklen, S., (1998). *Qualitative research for education: An introduction to theory and methods.* Needham Heights, MA: Allyn & Bacon.

Bolon, C. (2002). Response to Michelson and to Willson and Kellow. *Educational Policy Analysis Archives.* Available: (http://epaa.asu.edu/epaa/v8n42.html).

Bolon, C. (2000). School-based standard testing. *Education Policy Analysis Archives* Online. Available: http://epaa.asu.edu/epaa/vol8n23.html

Boyatzis, R. (1998). *Transforming qualitative information: Thematic analysis and code development.* Thousand Oaks, CA: Sage Publications.

Branch, A. (2000). New York schools continue to dig out from testing and attendance scandals. *Curriculum Administrator,* 36(2), pp. 8-9.

Camilli, G. (2000). Texas gains on NAEP: Points of light? Educational Policy Analysis Archives, 8 (42). Available: (http://epaa.asu.edu/epaa/v8n42.html).

College Board (2003). *About ETS.* Educational Testing Service. Available: http://www.ets.org/about/history/index.html

College Board (2003). *New SAT.* Educational Testing Service. Available: http://www.collegeboard.com/about/neewsat/history.html

Davies, A. (2000). *Making classroom assessment work.* Merville, British Columbia, Canada: Connections Publishing.

Education Assessment Insider (2002). Denville, NJ: Aspen Publishers, Inc. May 1 (2).

Eisner, E. (1998). *The enlightened eye: Qualitative inquiry and the enhancement of educational practice.* Upper Saddle River, NJ: Prentice Hall.

Fair Test: Testing our children (2003). National Center for Fair and Open Testing: Cambridge, Mass. Available: (http://fairtest.org/facts/csrtests.html).

Gallup, A. (2001). The 33rd annual Phi Delta Kappa poll of the public's attitudes toward the public schools. *Phi Delta Kappan,* September, 2002, 41-58.

Garrett, S. (2001). Oklahoma school testing contracts awarded by state board of education. Oklahoma State Department of Education Online. Available: http://sde.state.ok.us/TestVendor.html

Gehring, J. (2002). Achieve says Oklahoma sends 'mixed signal' on goals. *Education Week,* 9/11/2002, p 18.

Goals 2000. U.S. Department of Education. Available: http://ed.gov/pubs/

Gullatt, D. & Ritter M. (2002). What is your grade? *American School Board Journal*, Feb., p 40-42.

Haney, W. (2000). The myth of the Texas miracle in education. *Education Policy Analysis Archives*. Available: (http://epaa.asu.edu/epaa/v8n41).

Keating, F. (2000). Oklahoma State of the State Address, Feb.7, 2000. Available: http://www.odl.state.ok.us.oar/governors/addresses/keating2000.pdf.

Keller, B. (2002). Gov. Dean questions wisdom of accepting ESEA money. *Education Week*, May 1.

Klein, S. et al. (2000). What do test scores in Texas tell us? Education Policy Analysis Archives. Available: (http://olam.ed.asu.edu/epaa/v8n49).

Langenbach, M, Vaughn, C., & Aagaard, L. (1994). *An introduction to education research.* Boston: Allyn & Bacon.

Majchrzak, A. (1984). *Methods for policy research: Applied social research methods series*, Vol. 3. Newbury Park, CA: Sage Publications.

Marzano, R. (2003). *What works in schools: Translating research into action.* Alexandria, VA: Association for Supervision and Curriculum Development.

Mehrens, W. (1998) Consequence of assessment: What is the evidence? *Education Policy Analysis Archives* Online. Available: http://epaa.asu.edu/epaa/vol6n13.html

McNeil, L. (2000). *Contradictions of School Reform: Educational costs of standardized testing.* New York: Routledge.

Meier, D. (1995). *The Power of Their Ideas*. Boston: Beacon Press.

Meisels, S. (2003) Creating a system of accountability: The impact of instructional assessment on elementary children's achievement test scores. *Education Policy Analysis Archives*. Available: (http://epaa.asu.edu/epaa/vlln9).

Merriam, S. (1998). *Qualitative research and case study applications in education* (revised and expanded from *Case study research in education*, 1994). San Francisco: Jossey-Bass Publishers.

Minority students lag in every grade. (2000). *Curriculum Review,* 39(6), 4.

Moustakas, C. (1990). *Heuristic research: Design, methodology, and applications*. Newbury Park, CA: Sage Publications.

National Assessment of Educational Progress (2003). National Center for Education Statistics. Available: http://nces.ed.gov/nationsreportcard.about

No Child Left Behind Act (2001). U. S. Department of Education. Available: http://www.nochildleftbehind.gov

Neill, M. (2003). High stakes, high risk. *American School Board Journal,* 190 (2), 18-21.

Newkham, S. (2002). Evaluation and Testing Program at the University of Oklahoma. Interview, October, 2002.

Noyce, P., Perda, D. & Traver, R. (2000). Creating data-driven schools. *Educational Leadership*, 57(5), 52-56.

Oklahoma Academic Performance Index (2000-2001). Overview from Oklahoma Title 70 O.S. § 3-150. Oklahoma State Department of Education.

Oklahoma School Testing Program Act (2001). Oklahoma State Department of Education. Article V: School Laws of Oklahoma.

Oklahoma State Department of Education (2003). Student Assessment, Oklahoma School Program Testing Act. Available: (http://www.sde.state.ok.us).

Olson, L. (2001). Finding the right mix. Education Week, Quality Counts Special , January 11, 12-20.

Overview about MCAS (2003). *Eye on Education.* Available: (http://www.eyeoneducation).

Paige, R. (2002). An overview of America's education agenda. *Phi Delta Kappan,* May, 2002, 708-13.

Palmer, P. (1998). *The courage to teach.* San Francisco: Jossey-Bass, A Wiley Company.

Patton, M. (1987). *How to use qualitative methods in evaluation* . Newbury Park, CA: Sage Publications, Inc.

PBS Frontline (2003). The testing industry's big four. Available: http://pbs.org/wgbh/pages/frontlines/shows/schools/testing/companies.html.

Popham, J (2002). Right task, wrong tool. *American School Board Journal,* February, 188 (16), 19-22.

Popham, J (2003). Trouble with testing. *American School Board Journal,* 198(2), 14-17.

Priority Academic Student Skills (PASS), Revised May 2000. Oklahoma State Department of Education.

Riley, R. (2002). Education reform through standards and partnerships, 1993-2000. *Phi Delta Kappan,* May, 2002, 700-707.

Roschewski, P. (2003). Nebraska STARS line up. Phi Delta Kappan 84 (7), 517-20.

Sack, J. (2003). 'No child' law vies for scarce state resources. *Education Week,* January 8.

Sagor, R. (2000). Guiding school improvement with action research. Alexandria, VA: Association for Supervision and Curriculum Development.

School Laws of Oklahoma. (2002). Oklahoma State Department of Education. Available: http://www.sde.state.ok.us

Shaw, D. (2000). Data mining. *Curriculum Administrator*, 36(3), 54-57.

Sirotnik, K. (2002). Promoting responsible accountability in schools and education. *Phi Delta Kappan*, 83(9), 662-671.

Spending big bucks on standardized tests. (2000). *Curriculum Review*, 39(5), 2.

Stake, R. (1998). Some comments on assessment in U.S. education. *Education Policy Analysis Archives* Online. Available: http://epaa.asu.edu/epaa/vol6n14.html

STARS Reporting System (2003). Nebraska Department of Education. Available at: (http://www.nde.state.ne.us/stars/).

Stiggins, R. (2001A). *Student-involved classroom assessment* (3rd ed.). Upper Saddle River, NJ: Merrill-Prentice Hall.

Stiggins, R. (2001B) *The unfulfilled promise of classroom assessment.* Educational Measurements and Practice, Fall.

Toenjes, L. & Dworkin, A. (2002) Are Increasing test scores in Texas really a myth, or is Haney's myth a myth? *Education Policy Analysis Archives.* Available: (http:olam.ed.asu.edu/epaa/v10n17).

Tofig, D. (2003). Pressure of tests not just on kids. *The Atlanta Journal-Constitution,* April 14.

Vaishnav, A. (2003). Judge hears arguments on MCAS graduation rule. *The Boston Globe,* March 26.

Wallace, D. (2000). Results, results, results? *Educational Leadership*, 57(5), 66-68.

What the numbers say. (2000). *Curriculum Review*, 39(5), 2.

Zukas, T. (2000). So you want to implement reform. *Educational Leadership*, 57(7), 54-55.

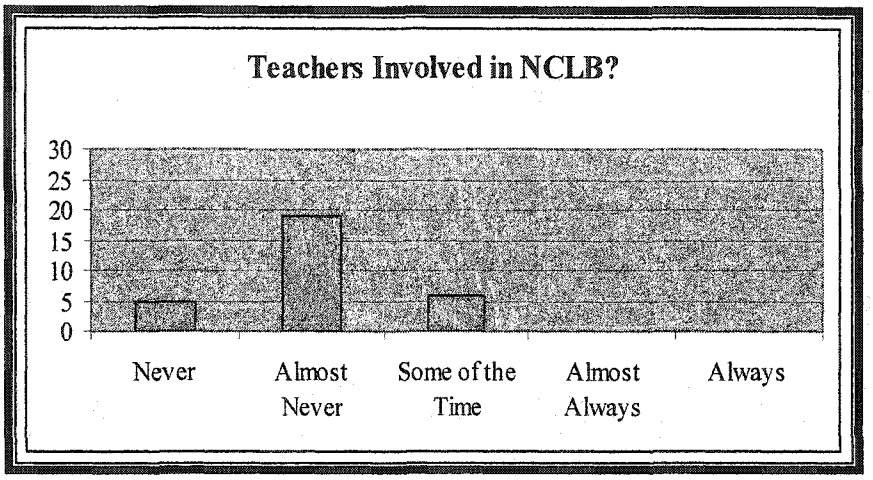
APPENDIX A

Survey Instrument and Results for all district Third and Fifth Grade Teachers

Check Only One Box 1=Never; 2=Almost Never; 3=Some of the Time; 4=Almost Always; and 5=Always					
	1	2	3	4	5
1. National leaders have involved teachers in determining goals and mandates for No Child Left Behind (NCLB)	5	19	6	0	0
2. National leaders have provided teachers with the resources to implement NCLB	9	15	6	0	0
3. Oklahoma leaders have involved teachers in determining the standards, or Priority Academic Student Skills (PASS)	2	6	16	6	0
4. Oklahoma leaders provided adequate, appropriate resources to classroom teachers for NCLB	6	9	15	0	0
5. Oklahoma has well-defined, clear standards in place for implementing NCLB	1	10	14	1	4
6. Oklahoma has appropriate assessments in place to track adequate yearly progress & student achievement for NCLB	1	10	14	5	0
7. Oklahoma provides test results to local districts in a time frame that allows for changes in instruction	1	5	15	9	0
8. Oklahoma provides student achievement results in reports that are easy to read and understand	1	5	13	11	0
9. My district has a well-defined, clear curriculum in place for implementing NCLB	0	0	7	11	12
10. My district has adequate, appropriate assessments in place to track progress for NCLB	0	0	5	15	10
11. My district regularly examines results of student performance from standardized achievement tests			5	7	18
12. My district regularly examines results of student performance from district teacher-made assessments			5	13	12
13. My district provides "time to learn" for all students who need extra time and help in reading and math			8	13	9
14. My district provides a "focus on learning" with every stakeholder involved in student achievement	2	2	14	6	6
15. My district promotes students' problem-solving and creative skills that are not measured by standardized tests.		2	3	13	12
16. As a classroom teacher, I plan instruction according to the Oklahoma PASS standards			3	8	19
17. As a teacher, I plan instruction according to the mandated Oklahoma achievement tests rather than the standards		2	10	10	8
18. As a teacher, I plan interventions according to the Oklahoma achievement test data		1	1	10	14
19. 4As a teacher, I plan interventions according to the district achievement test data			11	11	8
20. As a teacher, I regularly use my own assessments to judge student progress			3	10	17
As teachers in Oklahoma, the years of experience are a total of 428 years, Mean = 16.5, Median = 17 years					

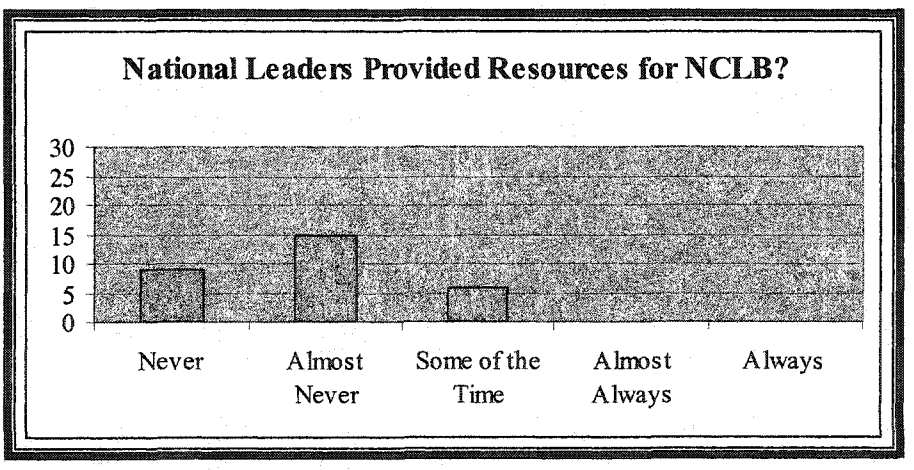
APPENDIX A 1 (Continued)

1. Have National Leaders Involved teachers in determining goals and mandates of NCLB? (5=Never; 19=Almost Never; 6= Some of the Time)



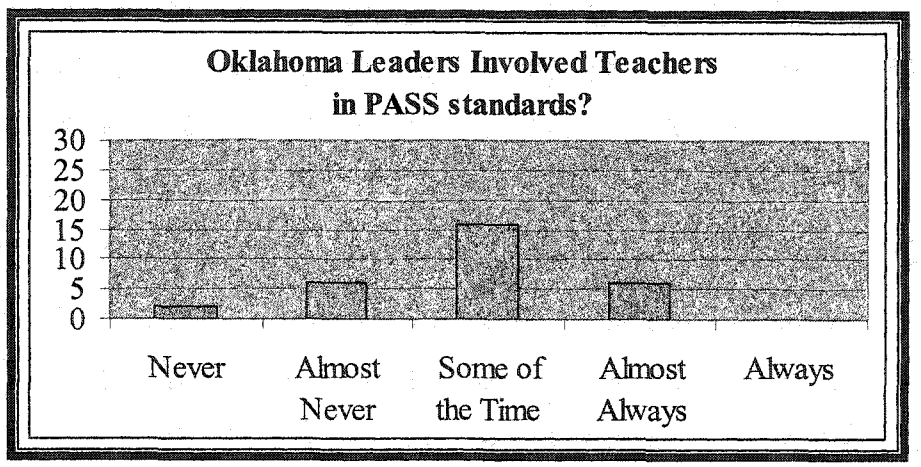
2. Have National leaders provided teachers with the resources to implement NCLB?

(9=Never; 15=Almost Never; 6=Some of the Time)

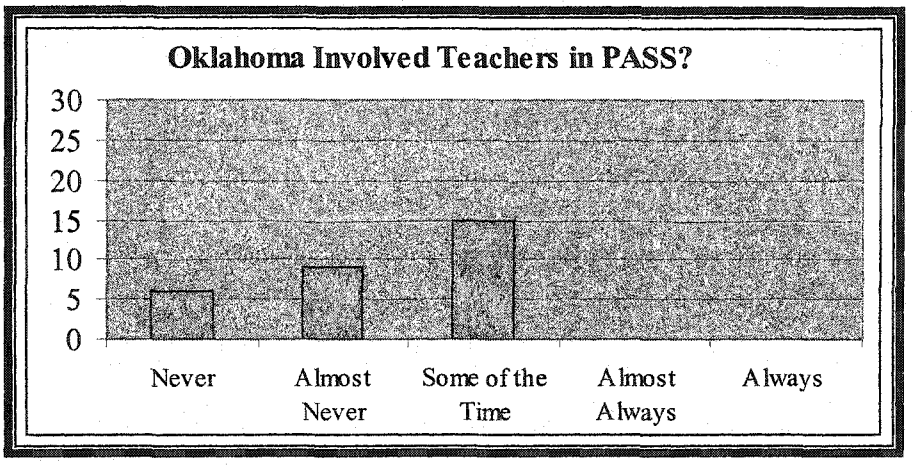


APPENDIX A 2 (Continued)

3. Have Oklahoma leaders involved teachers in determining the standards, the Priority Academic Student Skills (PASS)? (2=Never; 6=Almost Never; 16=Some of the Time; 6=Almost Always)

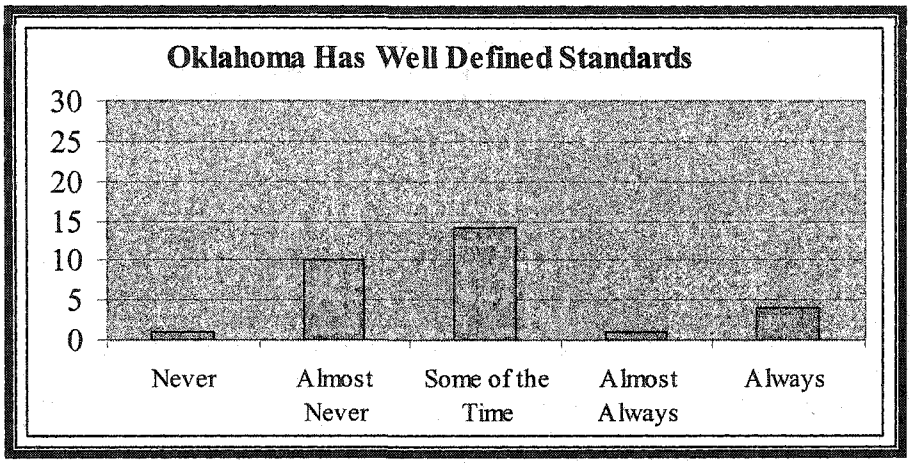


4. Have Oklahoma leaders provided adequate, appropriate resources to teachers for NCLB? (6=Never; 9=Almost Never; 15=Some of the Time)

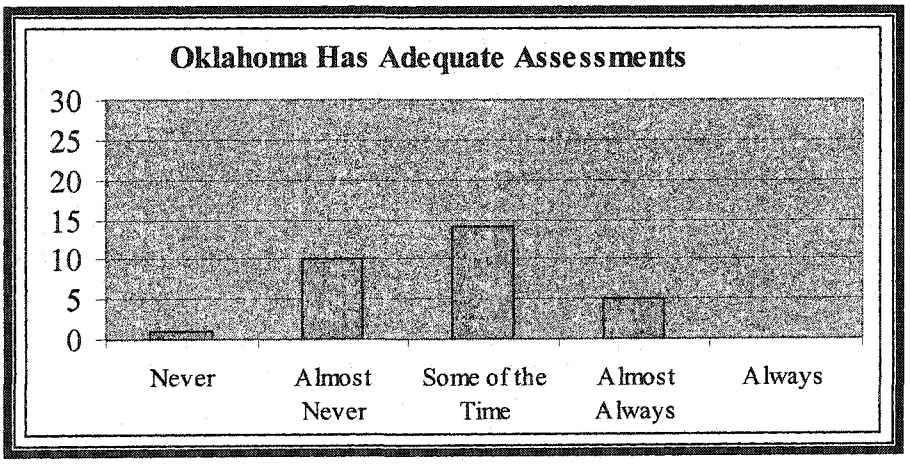


APPENDIX A 3 (Continued)

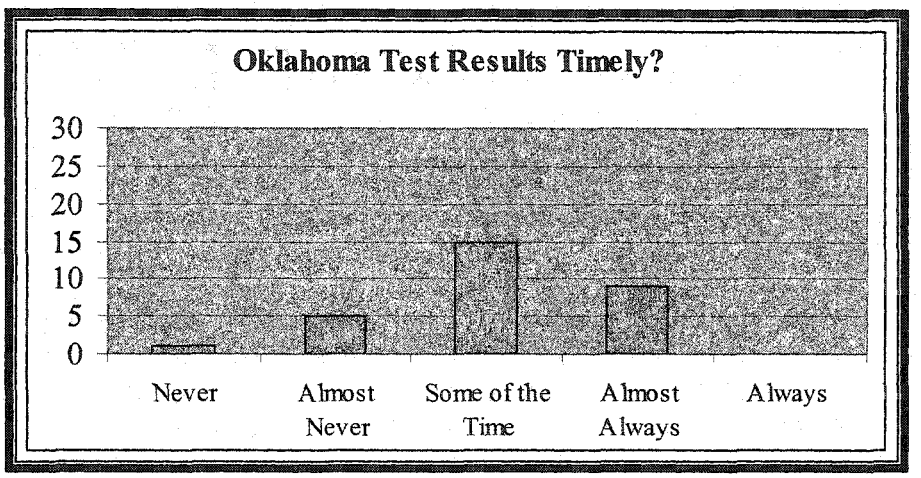
5. Does Oklahoma have well-defined, clear standards in place for implementing NCLB? (1=Never; 10=Almost Never; 14=Some of the Time; 1=Almost Always; 4=Always)



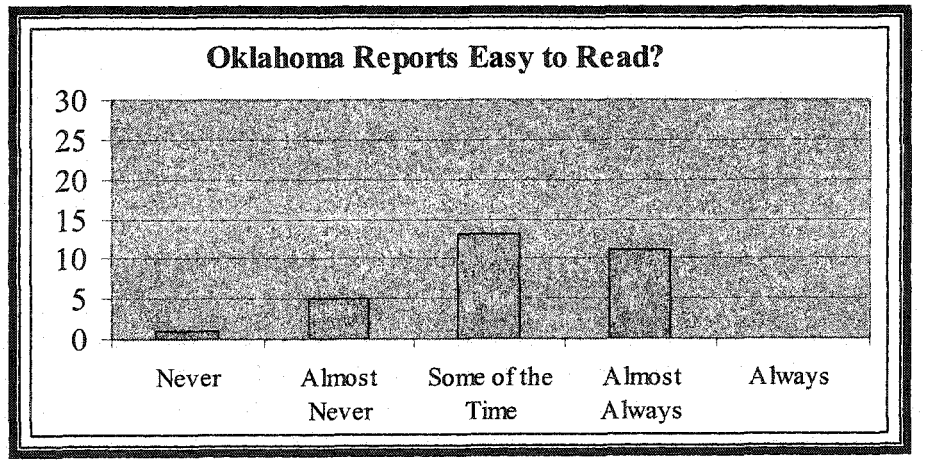
6. Does Oklahoma have adequate, appropriate assessments in place to track adequate yearly progress and student achievement for NCLB? (1=Never; 10=Almost Never; 14=Some of the Time; 5=Almost Always)



7. Has Oklahoma provided student achievement results to local districts in a timeframe that allows for changes in instruction? (1=Never; 5=Almost Never; 15=Some of the Time; 9=Almost Always)

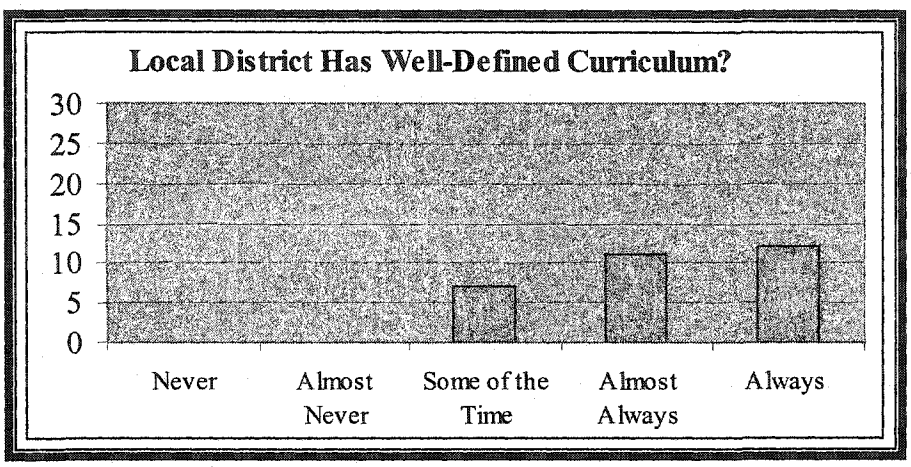


8. Does Oklahoma provide student achievement results in reports that are easy to read and understand? (1=Never; 5=Almost Never; 13=Some of the Time; 11=Almost Always)

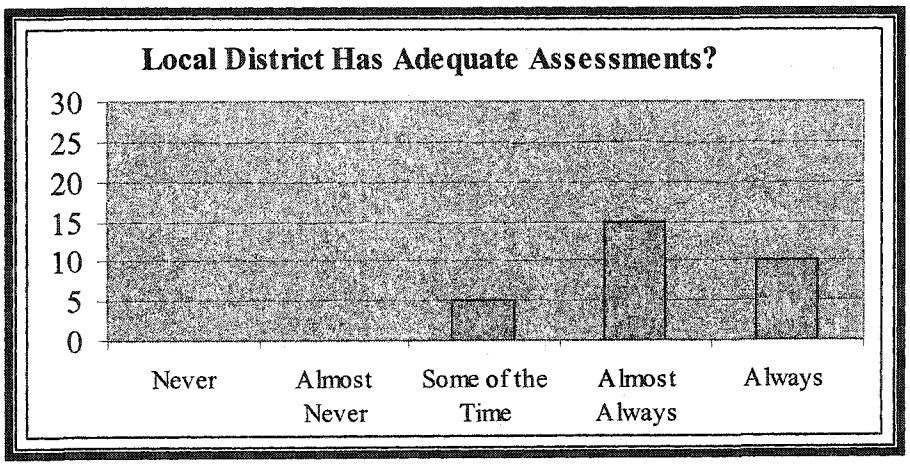


APPENDIX A 5 (Continued)

9. Does my district have a well-defined, clear curriculum in place for implementing NCLB? (7=Some of the Time; 11=Almost Always; 12=Always)

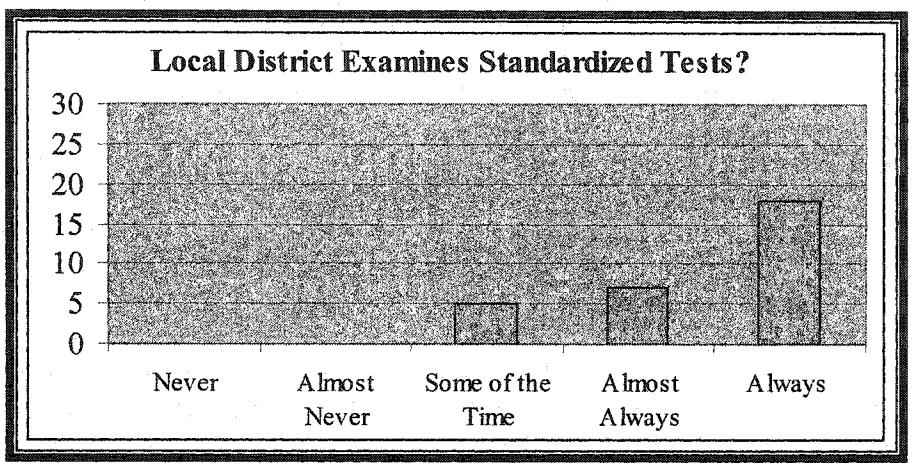


10. Does my local district have adequate, appropriate assessments in place to track progress for NCLB? (5=Some of the Time; 15=Almost Always; 10=Always)

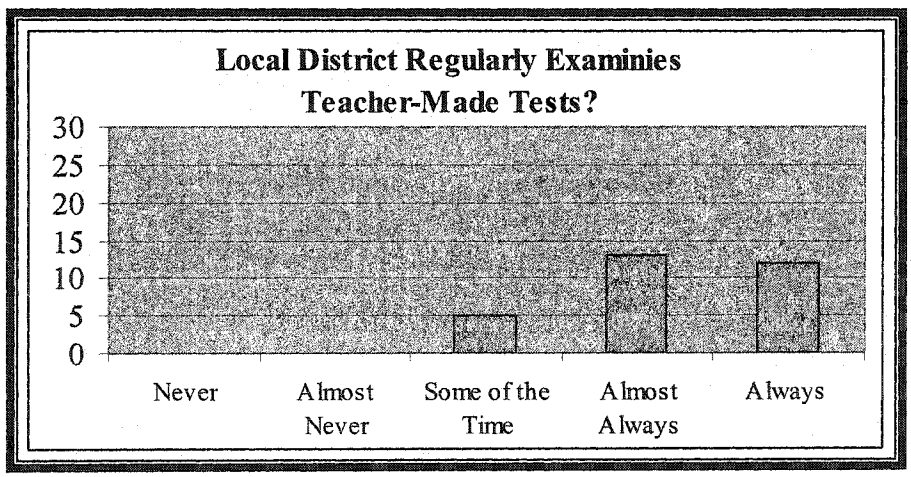


APPENDIX A 6 (Continued)

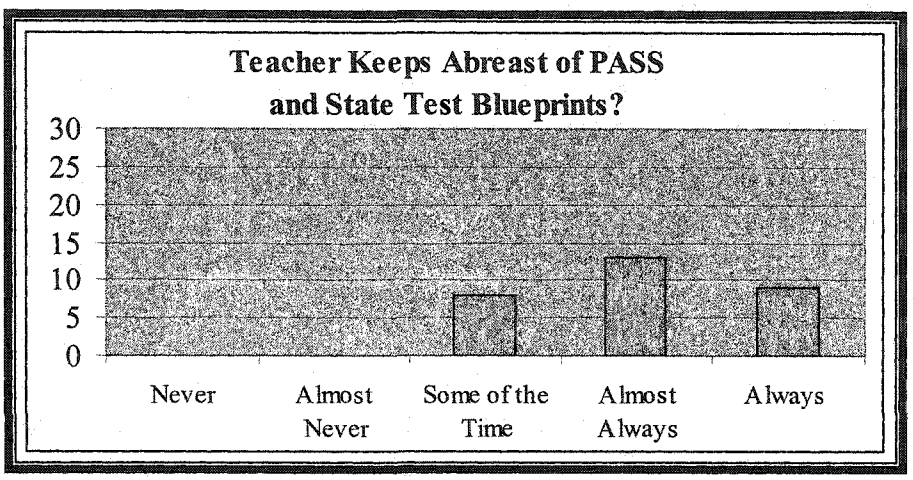
11. Does my local district regularly examine results of student performance from standardized achievement tests? (5=Some of the Time; 7=Almost Always; 18=Always)



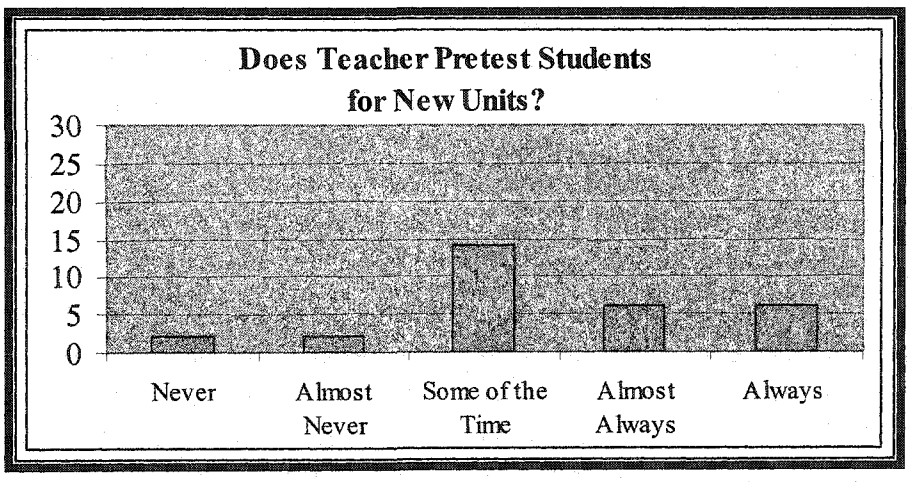
12. Does my district regularly examine results of student performance from district teacher-made assessments? (5=Some of the Time; 13=Almost Always; 12=Always)



13. As a classroom teacher, I keep abreast of content standards (PASS) and current state test blueprints? (8=Some of the Time; 13=Almost Always; 9=Always)

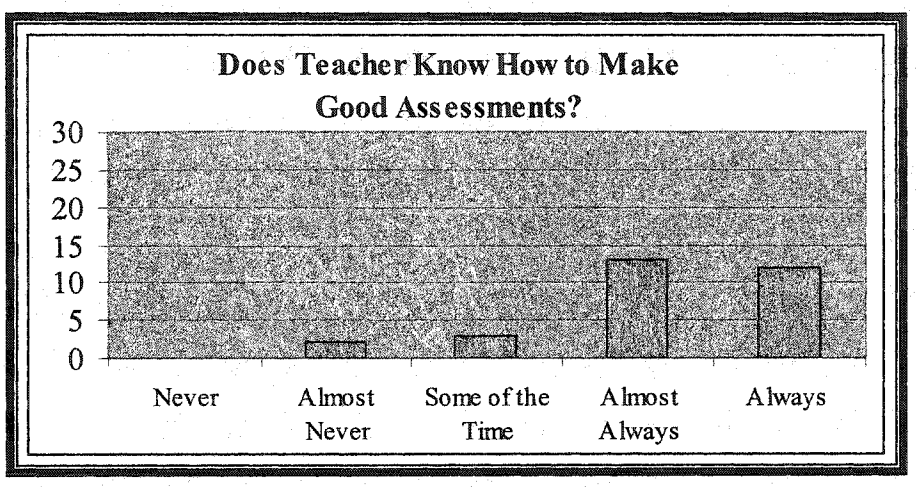


14. As a classroom teacher, I pretest my students before I start a new unit of study (e.g., fractions or punctuation). (2=Never; 2=Almost Never; 14=Some of the Time; 6=Almost Always; 6=Almost Always)

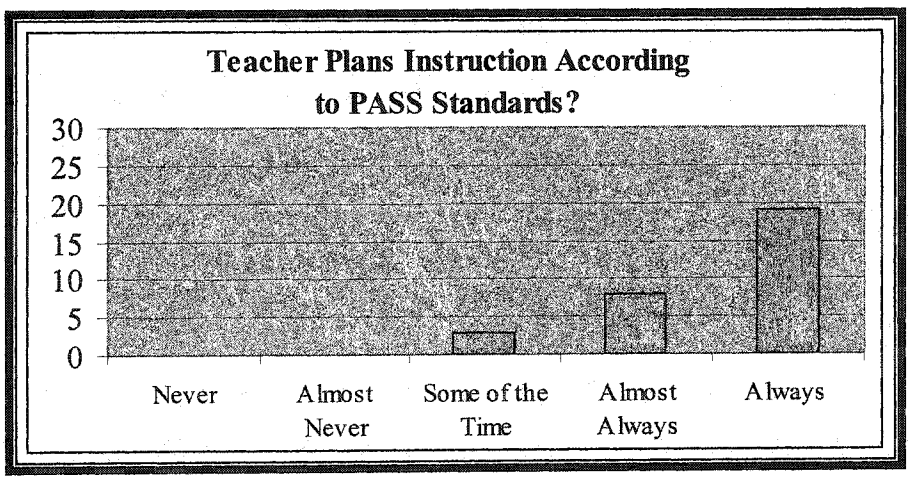


APPENDIX A 8 (Continued)

15. As a classroom teacher, I know how to build assessments that will give me good instructional information about my students. (2=Almost Never; 3=Some of the Time; 13=Almost Always; 12=Always)

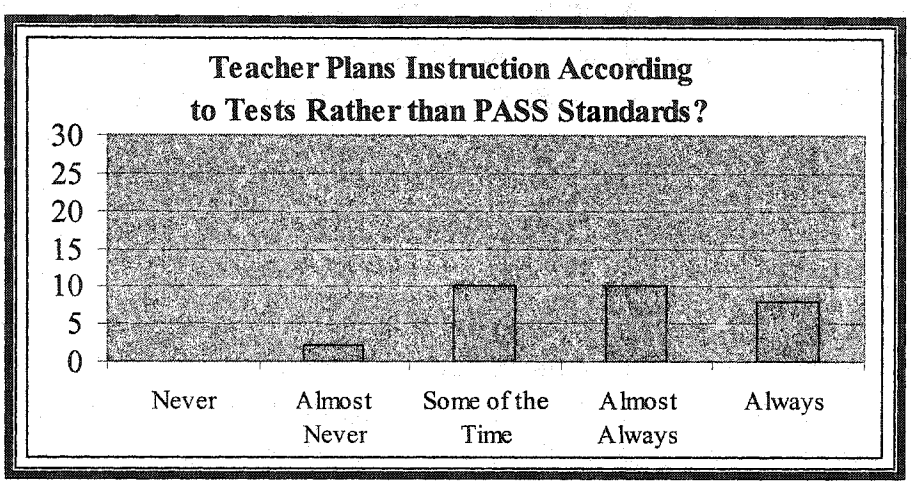


16. As a classroom teacher, I plan instruction according to the Oklahoma PASS standards? (3=Some of the Time; 8=Almost Always; 19=Always)

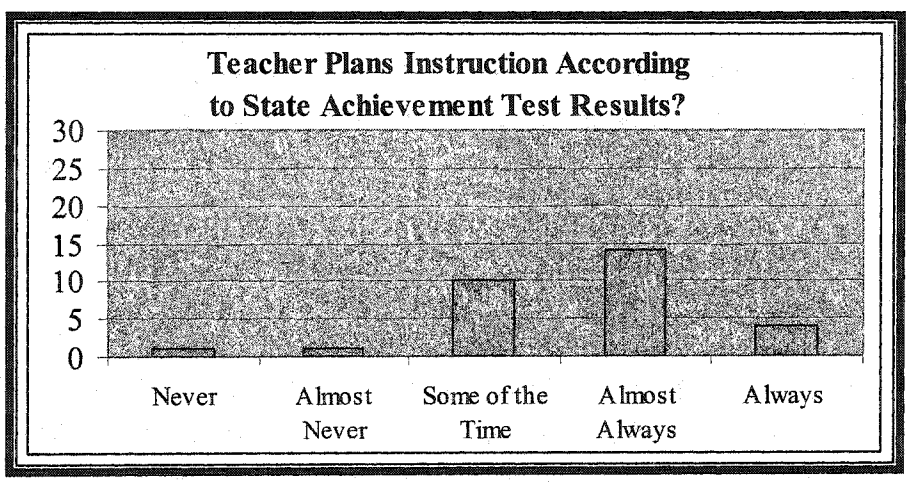


APPENDIX A 9 (Continued)

17. As a classroom teacher, I plan instruction according to the mandated Oklahoma achievement tests rather than the standards? (2=Almost Never; 10=Some of the Time; 10=Almost Always; 8=Always)

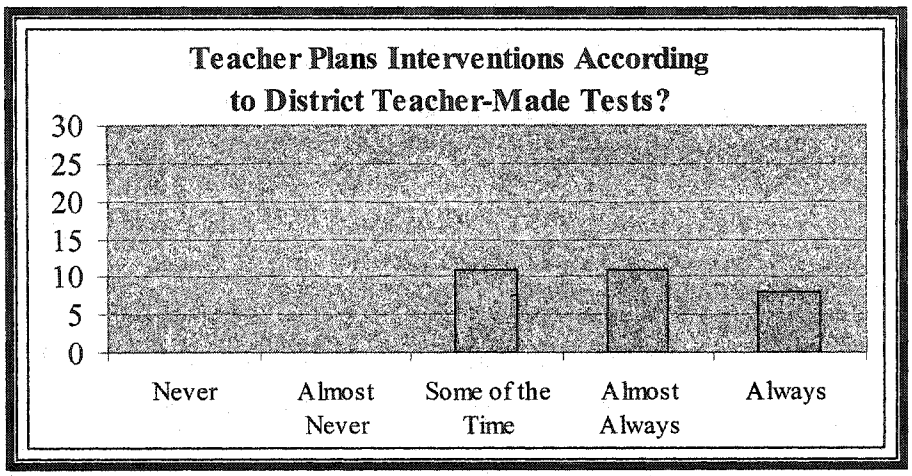


18. As a classroom teacher, I plan interventions according to the Oklahoma achievement data? (1=Never; 1=Almost Never; 10=Some of the Time; 14=Almost Always; 4=Always)

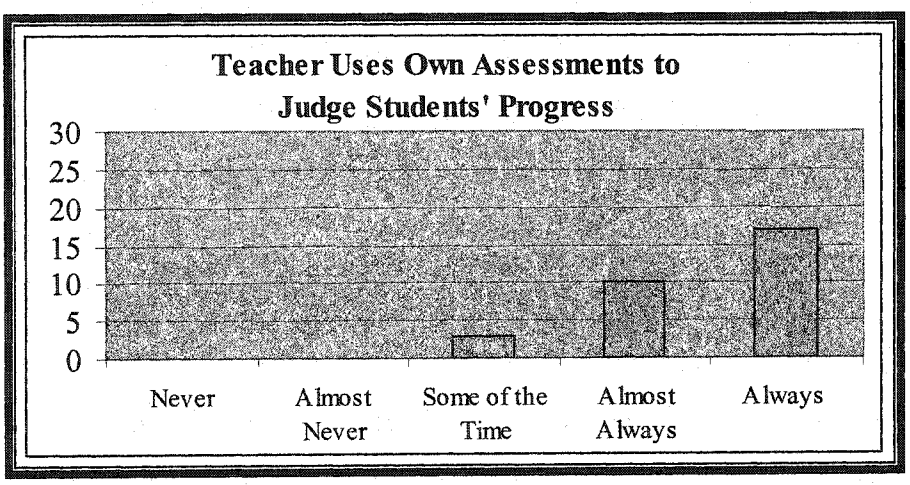


APPENDIX A 10 (Continued)

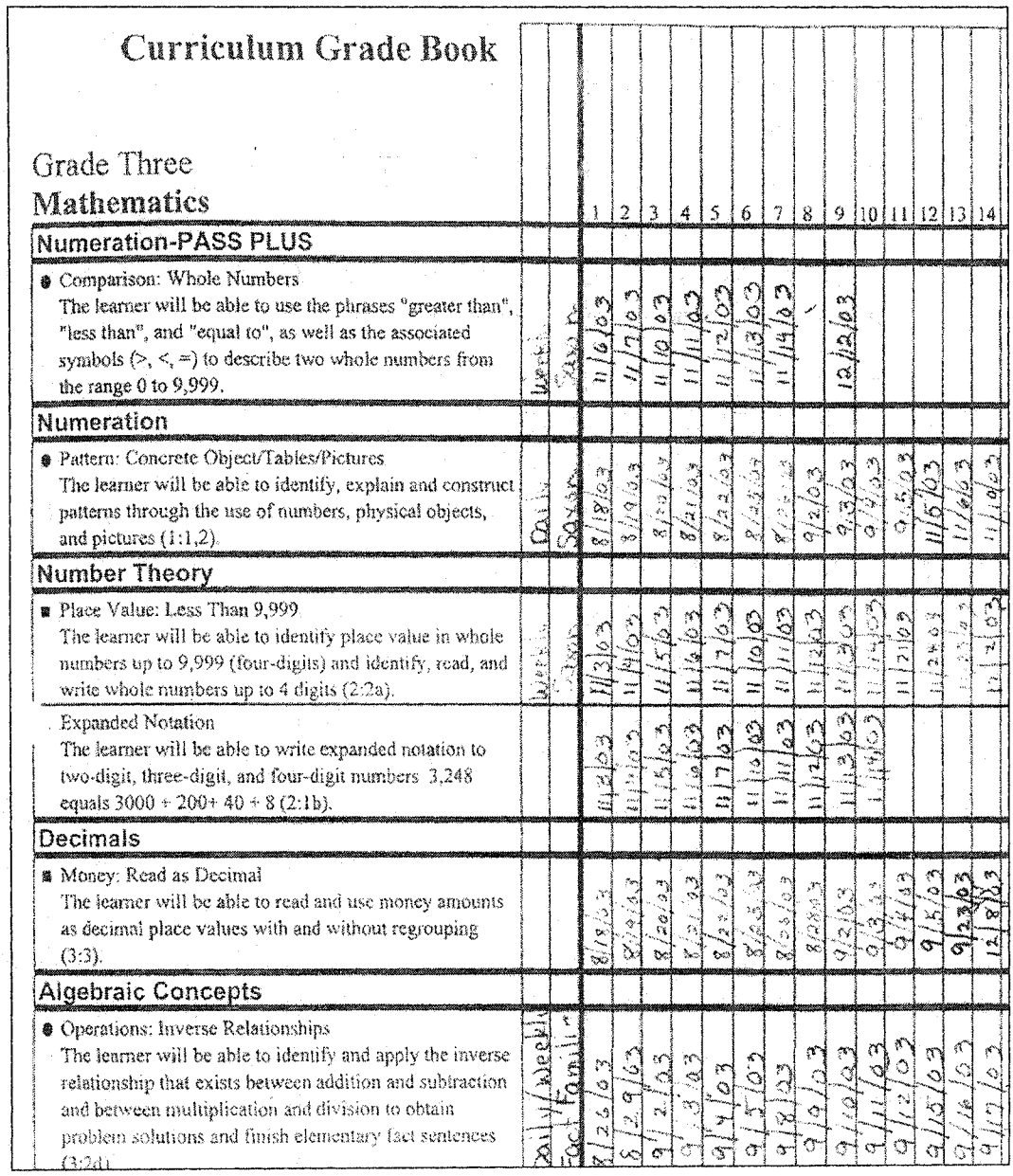
19. As a classroom teacher, I plan interventions according to the district achievement data (teacher-made tests). (11=Some of the Time; 11=Almost Always; 8=Always)



20. As a classroom teacher, I regularly use my own assessments to judge student progress? (3=Some of the Time; 10=Almost Always; 17=Always)



APPENDIX B
PASS Standards Documented by the Teacher



Curriculum Grade Book

Grade Three
Mathematics

			1	2	3	4	5	6	7	8	9	10	11	12	13	14
Numeration-PASS PLUS																
● Comparison: Whole Numbers The learner will be able to use the phrases "greater than", "less than", and "equal to", as well as the associated symbols (>, <, =) to describe two whole numbers from the range 0 to 9,999.	Work	Pass Pr	11/6/03	11/7/03	11/10/03	11/11/03	11/12/03	11/13/03	11/14/03		12/2/03					
Numeration																
● Pattern: Concrete Object/Tables/Pictures The learner will be able to identify, explain and construct patterns through the use of numbers, physical objects, and pictures (1:1,2).	Daily	SOAR	8/18/03	8/19/03	8/20/03	8/21/03	8/22/03	8/25/03	8/26/03	9/2/03	9/3/03	9/4/03	9/5/03	11/5/03	11/6/03	11/10/03
Number Theory																
■ Place Value: Less Than 9,999 The learner will be able to identify place value in whole numbers up to 9,999 (four-digits) and identify, read, and write whole numbers up to 4 digits (2:2a).	Work	Pass	11/3/03	11/4/03	11/5/03	11/6/03	11/7/03	11/10/03	11/11/03	11/12/03	11/13/03	11/14/03	11/21/03	11/24/03	12/1/03	
Expanded Notation The learner will be able to write expanded notation to two-digit, three-digit, and four-digit numbers 3,248 equals 3000 + 200+ 40 + 8 (2:1b).			11/3/03	11/4/03	11/5/03	11/6/03	11/7/03	11/10/03	11/11/03	11/12/03	11/13/03	11/14/03				
Decimals																
■ Money: Read as Decimal The learner will be able to read and use money amounts as decimal place values with and without regrouping (3:3).			8/19/03	8/19/03	8/20/03	8/21/03	8/22/03	8/25/03	8/26/03	8/27/03	9/2/03	9/3/03	9/4/03	9/5/03	9/23/03	12/8/03
Algebraic Concepts																
● Operations: Inverse Relationships The learner will be able to identify and apply the inverse relationship that exists between addition and subtraction and between multiplication and division to obtain problem solutions and finish elementary fact sentences (3:2d).	Daily/Weekly	Fact Families	8/26/03	8/29/03	9/2/03	9/3/03	9/4/03	9/5/03	9/8/03	9/10/03	9/10/03	9/11/03	9/12/03	9/15/03	9/16/03	9/17/03

Document/Artifact for
Table 8, #1, Teachers planned instruction according to Oklahoma PASS standards, documenting and dating the standards.

APPENDIX C
Oklahoma PASS Blueprint for Grade 5 Reading Teachers Used as A Guide

Priority Academic Student Skills

Oklahoma Core Curriculum Tests Blueprint
Beginning Spring 2003

Reading
Grade 5

PASS Clusters	Approximate Number of Items	Approximate Percentage of Test
Vocabulary (1.0)	**12**	**24%**
Words in Context (1.1)	4	
Affixes, Roots, and Stems (1.2)	4	
Synonyms, Antonyms, and Homonyms (1.3)	4	
Comprehension/Critical Literacy (3.0)	**21**	**42%**
Literal Understanding (3.1)	4	
Inferences and Interpretation (3.2)	5-7	
Summary and Generalization (3.3)	6-8	
Analysis and Evaluation (3.4)	6-8	
Literature (4.0)	**11**	**22%**
Literary Elements (4.2)	6-8	
Figurative Language/Sound Devices (4.3)	5-7	
Research and Information (5.0)	**6**	**12%**
Accessing Information (5.1)	6	
Total Test	***50***	***100%***

he test blueprint reflects the degree of representation given on the test to each *PASS* standard nd objective.

o access the current blueprint for each year with adjusted weights for standards and objectives, heck the Student Assessment Web site at <http://title3.sde.state.ok.us/studentassessment>.

Document/Artifact for
Table 8, #2, Teachers planned instruction according to Oklahoma PASS test blueprints.

APPENDIX D
Group Report for Third Grade Reading/Math/Language

Teachers Found This Document Difficult to Understand and Without Pertinent Information

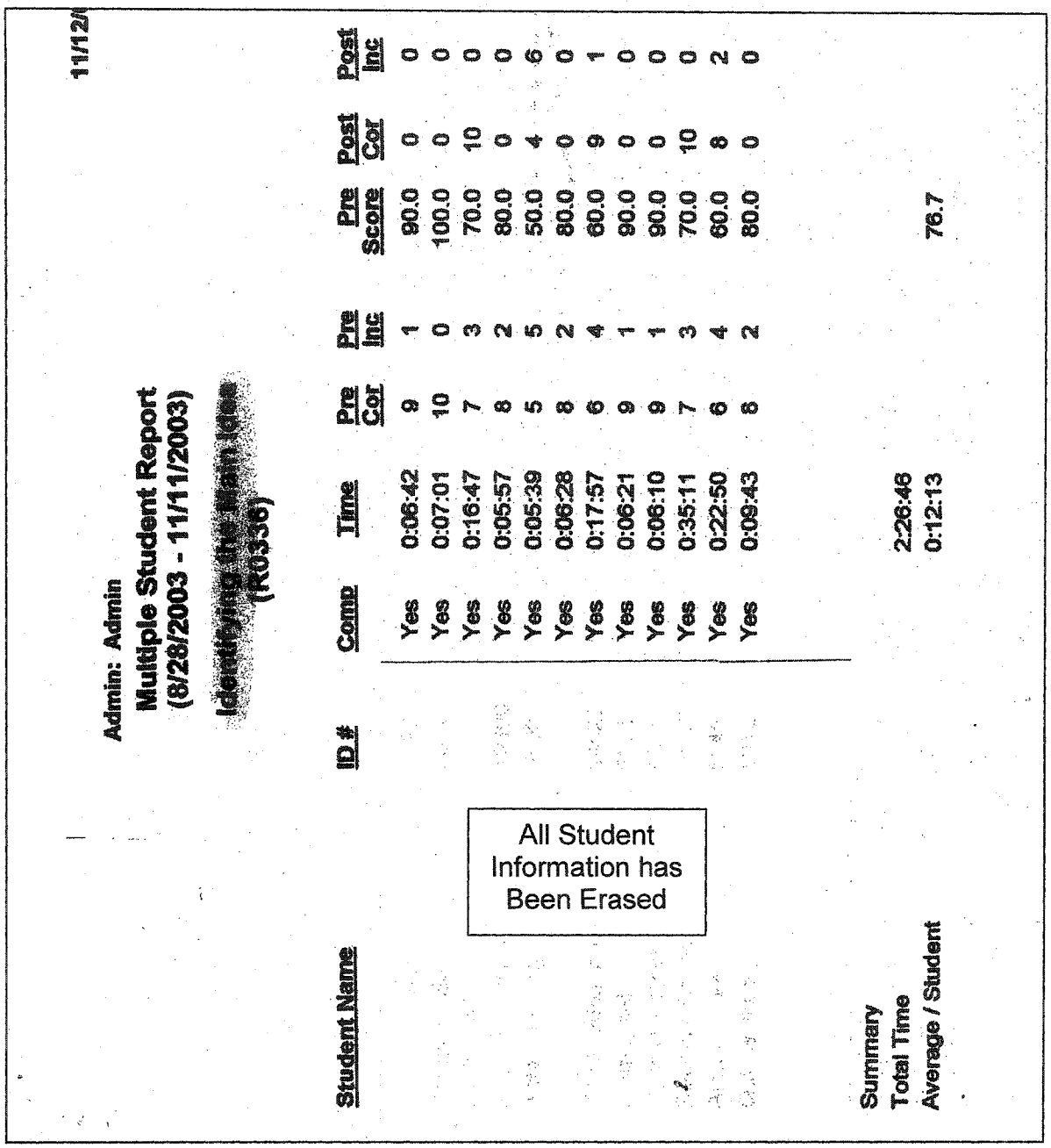
GRADE: 03
TEST DATE: 03/03

TEST TYPE: MULTIPLE CHOICE REGULAR STUDENTS

TOTAL NUMBER TESTED = 251	READING			MATHEMATICS			LANGUAGE	
	Total Reading	Reading Vocab	Reading Comp	Total Math	Problem Solving	Procedures	Language	
Number Possible	84	30	54	76	46	30	48	
Number Tested	247	247	249	249	249	249	231	
National Individual PR-S of Mean NCE	72-6	73-6	69-6	77-7	74-6	80-7	76-6	
Natl PR Summary								
76 - 99 N	103	118	99	133	120	146	122	
51 - 75 N	80	79	76	74	69	72	69	
26 - 50 N	47	42	56	34	49	19	29	
1 - 25 N	17	8	18	8	11	12	11	
76 - 99 %	42	48	40	53	48	59	53	
51 - 75 %	32	32	31	30	28	29	30	
26 - 50 %	19	17	22	14	20	8	13	
1 - 25 %	7	3	7	3	4	5	5	
Percent At/Above the National 50th PR	74	80	70	86	76	88	83	
National Stanine Summary								
(Above Avg) 7,8,9 N	103	90	99	133	120	146	122	
(Average) 4,5,6 N	133	149	137	108	119	94	100	
(Below Avg) 1,2,3 N	11	8	13	8	10	9	9	
(Above Avg) 7,8,9 %	42	36	40	53	48	59	53	
(Average) 4,5,6 %	54	60	55	43	48	38	43	
(Below Avg) 1,2,3 %	4	3	5	3	4	4	4	
Natl NCE Statistics								
Mean	62.4	63.0	60.5	65.3	63.3	67.4	64.8	
Standard Deviation	17.4	16.4	17.2	15.5	16.6	17.1	16.2	
Percentiles								
P90 - 90th	86.1	84.3	83.3	84.5	82.7	88.4	83.1	
Q3 - 75th	74.1	70.6	71.4	76.2	73.4	75.9	75.6	
Median - 50th	60.8	60.4	59.1	66.0	63.3	65.2	65.4	
Q1 - 25th	49.5	51.7	47.9	55.8	50.4	55.0	53.4	
P10 - 10th	41.0	41.2	36.4	44.9	42.4	45.5	43.6	

Document/Artifact for
Table 9, #3, State data should be more user-friendly, easier to read and understand

APPENDIX E
Diagnostic Computer Report for Multiple Students



11/12/

Admin: Admin

Multiple Student Report
(8/28/2003 - 11/11/2003)

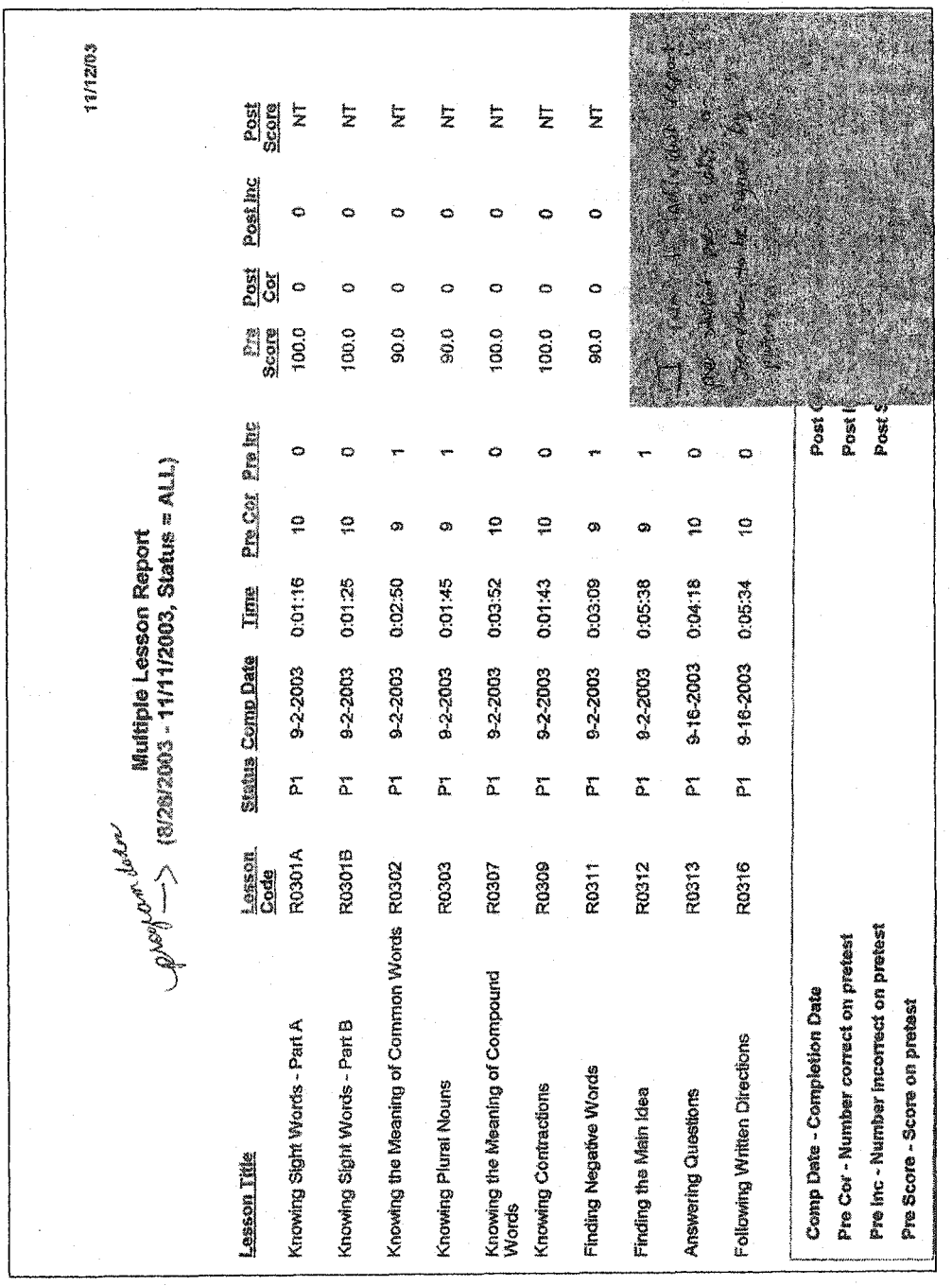
Identifying the Plain Idea
(R0336)

Student Name	ID #	Comp	Time	Pre Cor	Pre Inc	Pre Score	Post Cor	Post Inc
		Yes	0:06:42	9	1	90.0	0	0
		Yes	0:07:01	10	0	100.0	0	0
		Yes	0:16:47	7	3	70.0	10	0
		Yes	0:05:57	8	2	80.0	0	0
		Yes	0:05:39	5	5	50.0	4	6
		Yes	0:06:28	8	2	80.0	0	0
		Yes	0:17:57	6	4	60.0	9	1
		Yes	0:06:21	9	1	90.0	0	0
		Yes	0:06:10	9	1	90.0	0	0
		Yes	0:35:11	7	3	70.0	10	0
		Yes	0:22:50	6	4	60.0	8	2
		Yes	0:09:43	8	2	80.0	0	0
Summary			2:26:46			76.7		
Total Time			0:12:13					
Average / Student								

All Student Information has Been Erased

Document/Artifact for
Table 9, #7, Teachers used their own assessments and classroom data to plan for AYP.
Table 9, #9, Teacher-made data and assessments were an integral part of everyday instruction.
Table 10, #5, Diagnostic and placement assessments were an integral part of classroom instruction.
Table 11, #2, Teacher-made assessments gave valuable information for placement.

APPENDIX F
Multiple Lesson Report from Computer for One Student



11/12/03

Multiple Lesson Report

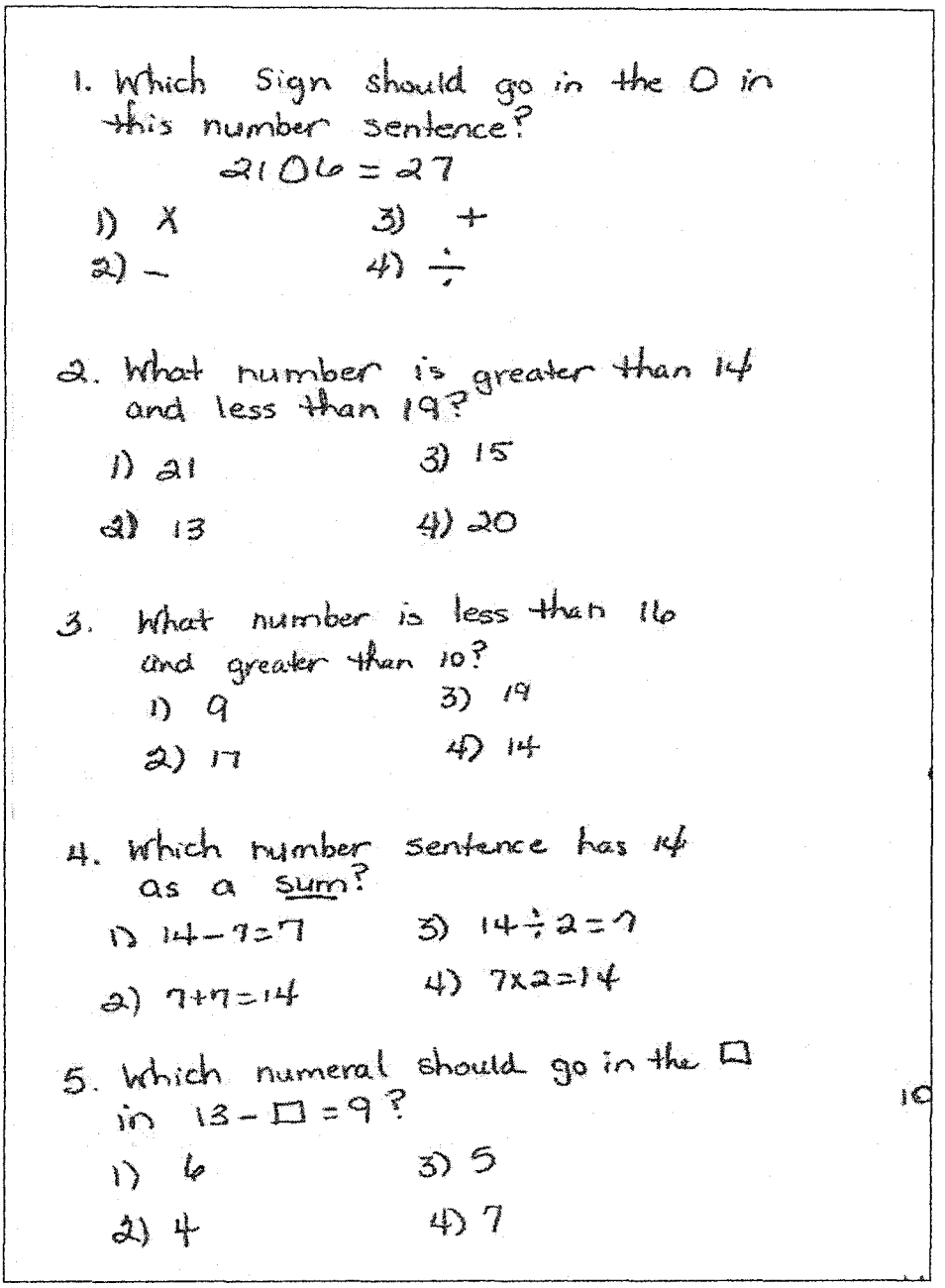
(8/26/2003 - 11/11/2003, Status = ALL)

Lesson Title	Lesson Code	Status	Comp Date	Time	Pre Cor	Pre Inc	Pre Score	Post Cor	Post Inc	Post Score
Knowing Sight Words - Part A	R0301A	P1	9-2-2003	0:01:16	10	0	100.0	0	0	NT
Knowing Sight Words - Part B	R0301B	P1	9-2-2003	0:01:25	10	0	100.0	0	0	NT
Knowing the Meaning of Common Words	R0302	P1	9-2-2003	0:02:50	9	1	90.0	0	0	NT
Knowing Plural Nouns	R0303	P1	9-2-2003	0:01:45	9	1	90.0	0	0	NT
Knowing the Meaning of Compound Words	R0307	P1	9-2-2003	0:03:52	10	0	100.0	0	0	NT
Knowing Contractions	R0309	P1	9-2-2003	0:01:43	10	0	100.0	0	0	NT
Finding Negative Words	R0311	P1	9-2-2003	0:03:09	9	1	90.0	0	0	NT
Finding the Main Idea	R0312	P1	9-2-2003	0:05:38	9	1				
Answering Questions	R0313	P1	9-16-2003	0:04:18	10	0				
Following Written Directions	R0316	P1	9-16-2003	0:05:34	10	0				

Comp Date - Completion Date
Pre Cor - Number correct on pretest
Pre Inc - Number incorrect on pretest
Pre Score - Score on pretest

Document/Artifact for
Table 9, #8, Teacher-made assessments were more helpful than state data to drive instruction;
Table 9, # 9 and Table 10, #5, Teacher-made data and assessments were an integral part of everyday instruction;
Table 10, #3, Use of data from classroom assessments on the current year's students was the most beneficial to teachers in showing students' actual progress.
Table 11, #2, Teacher-made assessments gave valuable information for placement.

APPENDIX G
Teacher Made Third Grade Assessment/Benchmark



1. Which sign should go in the O in this number sentence?

21 O 6 = 27

1) X 3) +
2) − 4) ÷

2. What number is greater than 14 and less than 19?

1) 21 3) 15
2) 13 4) 20

3. What number is less than 16 and greater than 10?

1) 9 3) 19
2) 17 4) 14

4. Which number sentence has 14 as a sum?

1) 14 − 7 = 7 3) 14 ÷ 2 = 7
2) 7 + 7 = 14 4) 7 x 2 = 14

5. Which numeral should go in the □ in 13 − □ = 9?

1) 6 3) 5
2) 4 4) 7

10

Document/Artifact for
Table 10, #4, Teachers created their own assessments to show progress and benchmark student growth for NCLB and AYP.

APPENDIX H
Computer-Generated Item Analysis for Reading

Reading Item Analysis		
Dictionary (guide words)	24%	**In this item analysis, the teacher plotted skills from a reading pre-test given to students at the beginning of the year. From this pretest, she identified the skills that few of the students knew (24%) and those that most (82%) of the students knew.**
Antonyms	35%	
Homonyms	35%	
Index	44%	
Suffix – Base Words	50%	
Base Words	56%	
Glossary	56%	**This teacher started her year in Reading with the skills where students needed most improvement. Then she gave them lots of practice during the year.**
Sight Words	59%	
Comprehension (fiction)	62%	
Table of Contents	68%	
Synonyms	76%	
Prefixes – Base Words	76%	
Comprehension (nonfiction)	82%	

Document/Artifact for
Table 11, #1 Item analyses from district or teacher-made pre tests were used.

APPENDIX I
Teacher-Made Assessment by Skill

Math, 5th Grade
1/04

Read each question carefully and notate the correct answer.

1. 7.1891
   x 0.211

   A. 1.5169001
   B. 1.43782
   C. 0.15169001
   D. 0.143782

2. What does the digit 1 mean in 0.131?

   A. hundredths
   B. tenths
   C. tens
   D. ones

3.. Reduce answer to lowest terms.

$$5\frac{2}{9} \div 1\frac{4}{8} =$$

   A. 5 4/9
   B. 7 5/6
   C. 5 1/9
   D. 3 13/27

4. Reduce answer to lowest terms.

$$8\frac{11}{12} \times 11\frac{3}{4} =$$

   A. 88 11/16
   B. 107/188
   C. 104 37/48
   D. 20 2/3

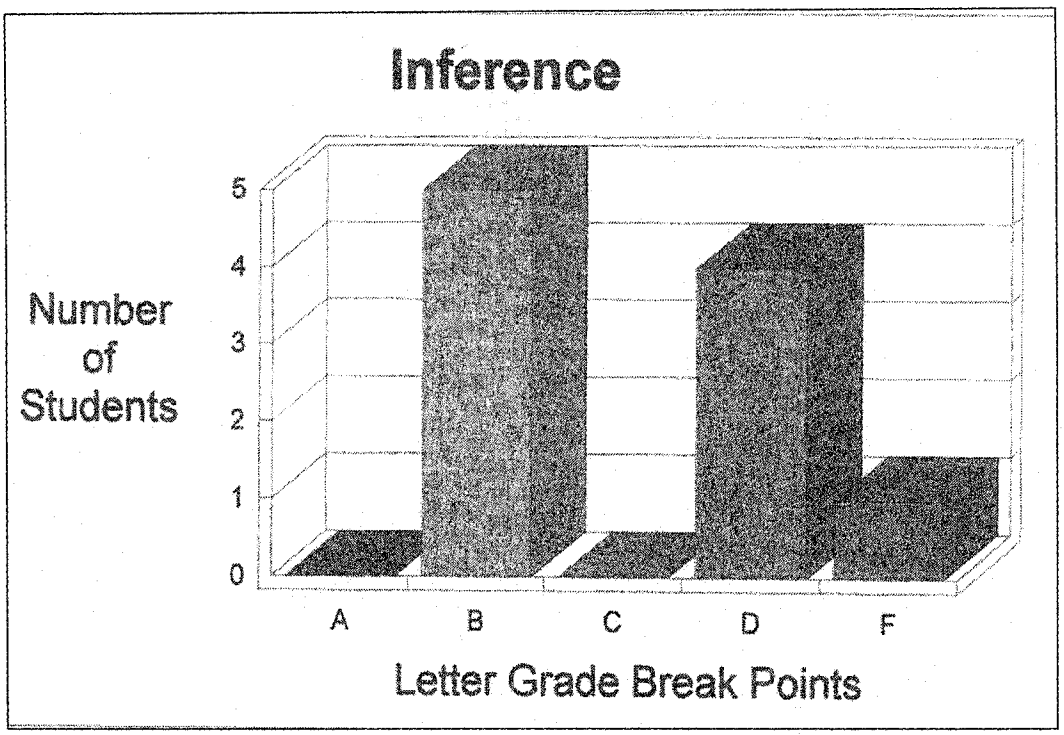
Math, 5th Grade

1. Multiply Decimals: Ten Thousandths
2. Place Value: Decimals - A
3. Divide Fractions: Mixed Numbers
4. Fractions for Part of a Set - B
5. Multiply Fractions: Mixed Numbers
6. Lines/ Segments/ Circles
7. Area of Rectangle - B
8. Mass/ Capacity - B
9. Volume - C
10. Ratio/ Proportion - A

Document/Artifact for
Table 11, #2, Teacher-made assessments gave valuable information for placement.

APPENDIX J
Chart of Student Progress by Letter Grade

Students Who Did Not Make a "B" (80%) Did the Assignments Again



Document/Artifact for
Table 11, #8, Graphs, charts, and timelines supplied instructional tools.
Table 14, #1, Effective teachers had high expectations of students.

APPENDIX K
Computer-Generated Lesson with Study Guide

STUDY GUIDE, MATH, GRADE 5

Decimals and Equivalent Fractions
Decimal numbers can be written in fraction format.

For example, 0.5 = 1/2.

It may be beneficial to concentrate on the student's understanding of decimal numbers before introducing fractions. Do not introduce fractions until decimals are completely understood. Once the student has a complete understanding of decimals, begin to use equivalent fractions. The following is a step-by-step example of how to calculate equivalent fractions for decimals.

Find the equivalent fraction for 0.75.

(1) (2)

$$.75 = \frac{75}{100} \qquad \frac{75}{100} = \frac{3}{4}$$

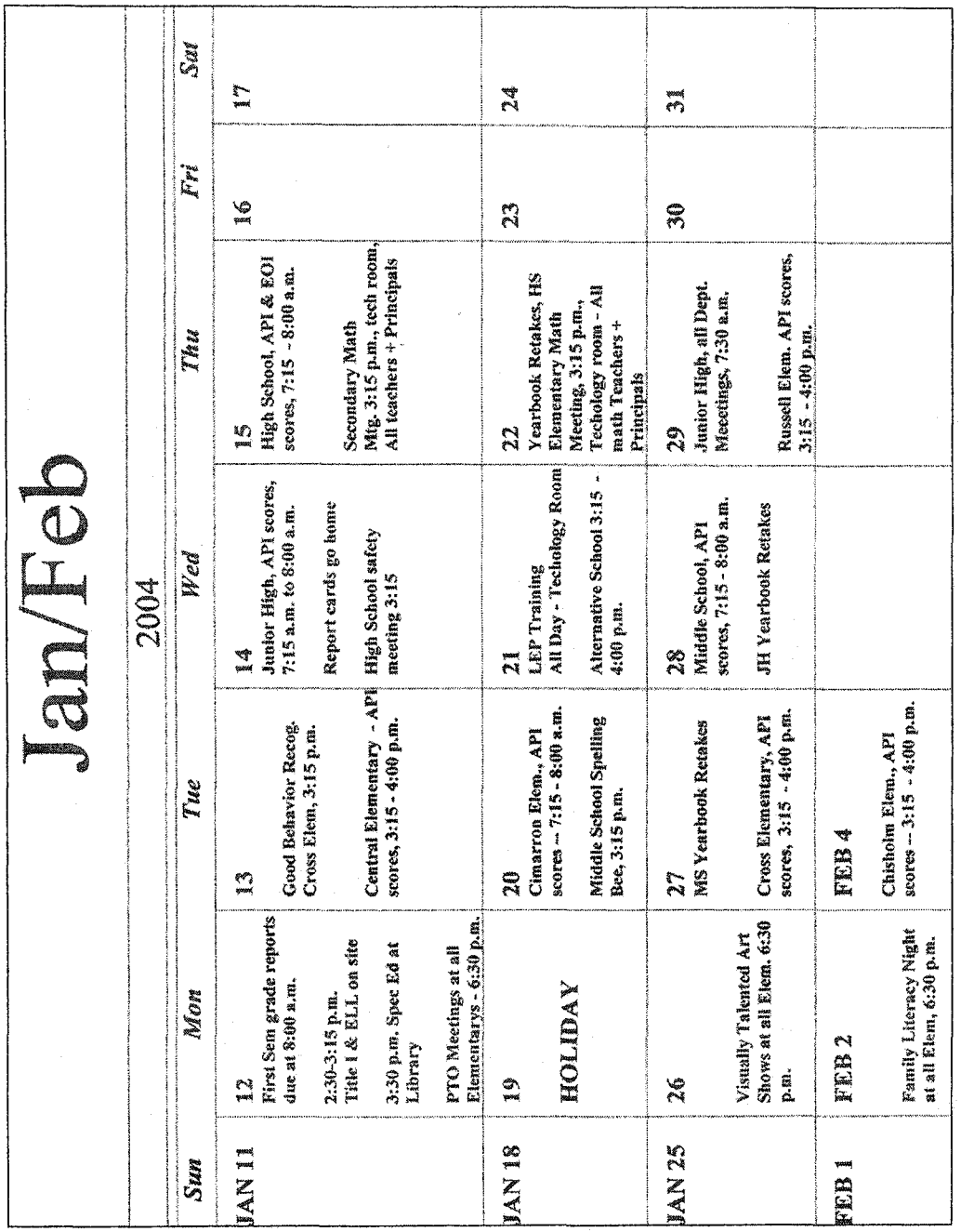
Step 1: Decimals are used to name tenths, hundredths, etc.
.75 is equal to 75 hundredths.
Convert .75 to a fraction of hundredths.

Step 2: Reduce the fraction to lowest terms. A fraction is in lowest terms when the numerator and denominator do not have a common factor greater than one. A common factor of 75 and 100 is 25.

Answer: 0.75 is equal to 3/4.

Document/Artifact for
Table 11, #10, Computer software supplemented classroom materials

APPENDIX L
Calendar of After-School Duties and Meetings



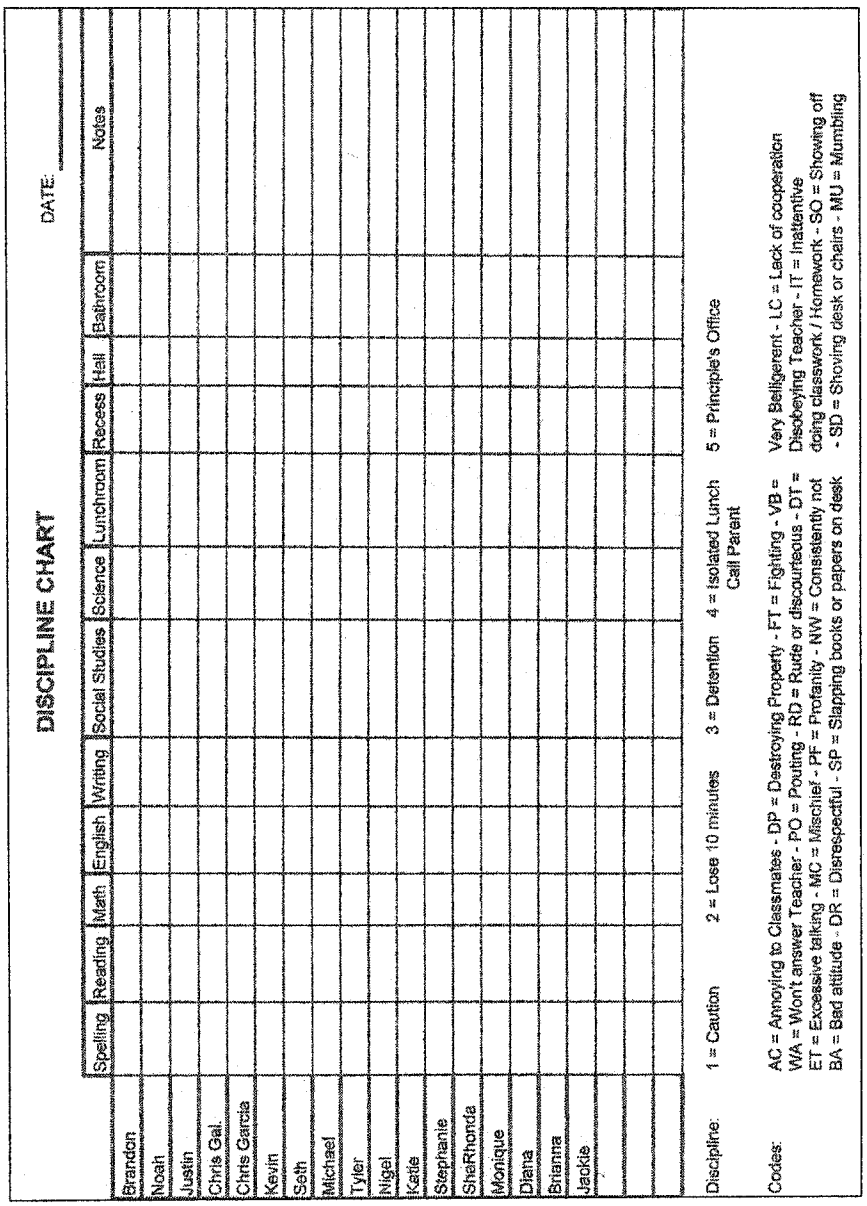
Jan/Feb

2004

Sun	*Mon*	*Tue*	*Wed*	*Thu*	*Fri*	*Sat*
JAN 11	12 First Sem grade reports due at 8:00 a.m. 2:30-3:15 p.m. Title 1 & ELL on site 3:30 p.m. Spec Ed at Library PTO Meetings at all Elementarys - 6:30 p.m.	13 Good Behavior Recog. Cross Elem, 3:15 p.m. Central Elementary - API scores, 3:15 - 4:00 p.m.	14 Junior High, API scores, 7:15 a.m. to 8:00 a.m. Report cards go home High School safety meeting 3:15	15 High School, API & EOI scores, 7:15 - 8:00 a.m. Secondary Math Mtg. 3:15 p.m., tech room, All teachers + Principals	16	17
JAN 18	19 HOLIDAY	20 Cimarron Elem., API scores -- 7:15 - 8:00 a.m. Middle School Spelling Bee, 3:15 p.m.	21 LEP Training All Day - Techology Room Alternative School 3:15 - 4:00 p.m.	22 Yearbook Retakes, HS Elementary Math Meeting, 3:15 p.m., Techology room - All math Teachers + Principals	23	24
JAN 25	26 Visually Talented Art Shows at all Elem. 6:30 p.m.	27 MS Yearbook Retakes Cross Elementary, API scores, 3:15 - 4:00 p.m.	28 Middle School, API scores, 7:15 - 8:00 a.m. JH Yearbook Retakes	29 Junior High, all Dept. Meeetings, 7:30 a.m. Russell Elem. API scores, 3:15 - 4:00 p.m.	30	31
FEB 1	FEB 2 Family Literacy Night at all Elem, 6:30 p.m.	FEB 4 Chisholm Elem., API scores -- 3:15 - 4:00 p.m.				

Document/Artifact for
Table 12, #4, After-school duties and meetings took additional "time" from preparation.

APPENDIX M
Teacher Documentation (Blank Form) of Discipline Events



DISCIPLINE CHART

DATE: _____

	Spelling	Reading	Math	English	Writing	Social Studies	Science	Lunchroom	Recess	Hall	Bathroom	Notes
Brandon												
Noah												
Justin												
Chris Gal.												
Chris Garcia												
Kevin												
Seth												
Michael												
Tyler												
Nigel												
Katie												
Stephanie												
SheRhonda												
Monique												
Diana												
Brianna												
Jackie												

Discipline: 1 = Caution 2 = Lose 10 minutes 3 = Detention 4 = Isolated Lunch Call Parent 5 = Principle's Office

Codes: AC = Annoying to Classmates - DP = Destroying Property - FT = Fighting - VB = Very Belligerent - LC = Lack of cooperation
WA = Won't answer Teacher - PO = Pouting - RD = Rude or discourteous - DT = Disobeying Teacher - IT = Inattentive
ET = Excessive talking - MC = Mischief - PF = Profanity - NW = Consistently not doing classwork / Homework - SO = Showing off
BA = Bad attitude - DR = Disrespectful - SP = Slapping books or papers on desk - SD = Shoving desk or chairs - MU = Mumbling

Document/Artifact for
Table 13, #1, Discipline was a major key in the effective teaching/learning process.

APPENDIX N
Documentation of Parent Involvement

Dear Parents,

I wanted each of you to know that I am thoroughly enjoying having your child in my classroom this year.

This year is just zooming by! In fact, next week your child will be receiving their first third grade report card.

To facilitate better parent/teacher communications, I will be starting telephone conferences with parents on a scheduled basis beginning Monday, October 19th. These conferences will be voluntary in nature and are for those parents who desire additional information concerning their child's progress.

For your convenience in scheduling, please provide the information requested below.

☐ I do not wish to be called for a scheduled telephone conference.

☐ I would like to be contacted by telephone every: _____ Two weeks

_____ Three weeks

Choose the day(s) most convenient. (First and second choice)

_____ Monday _____ Tuesday _____ Wednesday _____ Thursday

Chose the most convenient time.

_____ 4 – 5PM _____ 7 – 8PM _____ 8 – 9PM _____ 9 – 10PM

Have a great year.

Sincerely,

Document/Artifact for
Table 13, #1: Discipline was a major key in the effective teaching/learning process.
Table 13, #2: Effective teachers make an intense effort to involve parents in school.

APPENDIX N (Continued)
Documentation of Parent Involvement

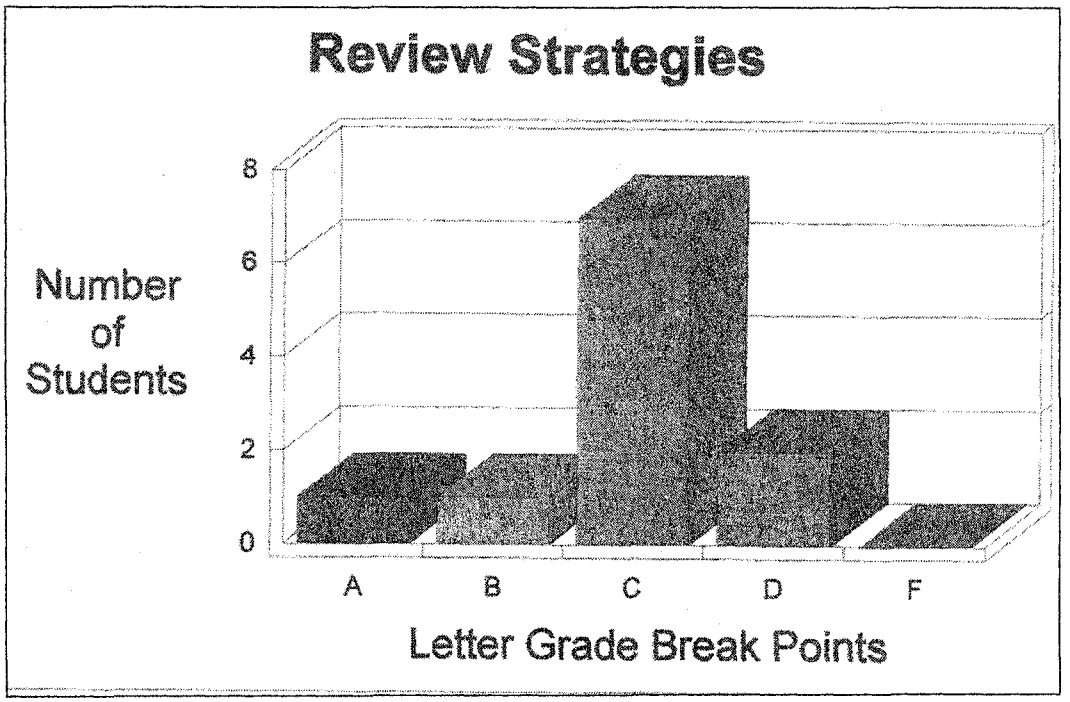
Parental Contact Log

ATE	TIME	CHILD	PARENT or GUARDIAN	COMMENTS

Document/Artifact for
Table 13, #1: Discipline was a major key in the effective teaching/learning process.
Table 13, #2: Effective teachers make an intense effort to involve parents in school.

APPENDIX O
Teacher Document of High Expectations

Chart of Student Progress by Letter Grade
Students Who Did Not Make a "B" (80%) Did the Assignment Again



Document/Artifact for
Table 11, #8, Graphs, charts, and timelines supplied instructional tools.
Table 14, #1, Effective teachers had high expectations of students.

Appendix P
Research Questions Used in Interviews

The following represent questions that were asked of six "highly effective" teachers in one southwest Oklahoma district, teachers who met "highly qualified" mandates under NCLB and who were also very successful and effective in the teaching/learning process as documented by students' test scores over the past five years.

1. How had "highly effective" teachers used state assessment data in planning instruction? What state data were actually used in planning instruction and how were that data used? What other data were used in the teaching/learning process?
2. How did "highly effective" teachers know what students were supposed to learn for their grade levels? What gave teachers that information? How did teachers know when students had learned what they were supposed to have learned?
3. How aligned were the skills assessed on Oklahoma state-mandated tests with PASS for both third grade NRT and fifth grade CRT? How did the teachers' delivered instruction align with the standards and the actual test?
4. How had "highly effective" teachers disaggregated test data into a useful form and had they been trained for that? When did they have time to analyze the data and when would be the most useful time to examine the year's instruction?
5. How had "highly effective" teachers decided what interventions were needed, how those were evaluated, and from what specific data were the decisions made.
6. How did teachers benchmark progress throughout the year?
7. Did student discipline, parental involvement, and student expectations impact learning in the classroom?