

DISSERTATION

HIGH-STAKES TESTING: THE STUDENT EXPERIENCE

Submitted by

Timothy Pearson

School of Education

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Spring 2009

UMI Number: 3374611

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[®]

UMI Microform 3374611
Copyright 2009 by ProQuest LLC
All rights reserved. This microform edition is protected against
unauthorized copying under Title 17, United States Code.

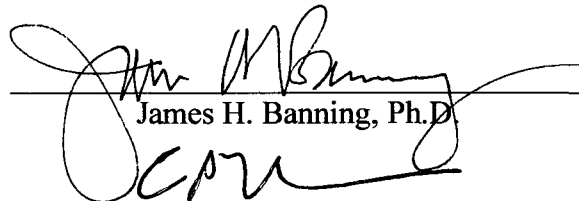
ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

COLORADO STATE UNIVERSITY

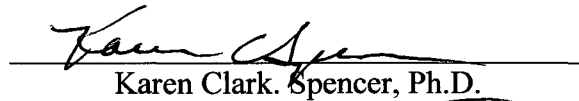
December 2, 2008

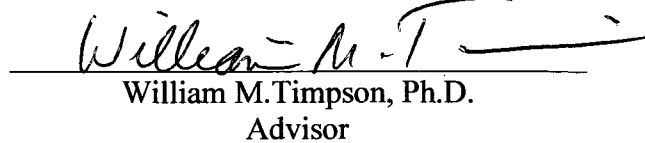
WE HEREBY RECOMMEND THAT THE DISSERTATION PREPARED UNDER OUR SUPERVISION BY TIMOTHY M. PEARSON ENTITLED HIGH-STAKES TESTING: THE STUDENT EXPERIENCE BE ACCEPTED AS FULFILLING IN PART REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY.

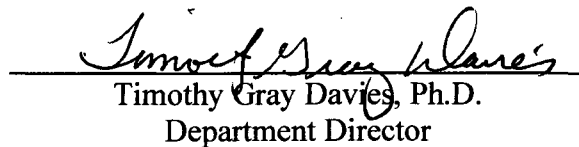
Committee on Graduate work


James H. Banning, Ph.D.


Clifford P. Harbour, J.D., Ed.D.


Karen Clark. Spencer, Ph.D.


William M. Timpson, Ph.D.
Advisor


Timothy Gray Davies, Ph.D.
Department Director

ABSTRACT OF DISSERTATION

HIGH-STAKES TESTING: THE STUDENT EXPERIENCE

The accountability components of the No Child Left Behind act of 2002 led to high-stakes testing in public schools in the United States. There are a large number of studies and articles related to this testing but little in the literature about how students perceive and understand high-stakes testing.

The purpose of this study was to investigate how students experience high-stakes tests. Six tenth grade students were interviewed for this narrative inquiry. Six narratives were constructed using transcript data. Six thematic elements were identified after additional analysis. The first thematic element is School Loyalty and Test Compliance. The participants varying levels of loyalty did not seem to affect their test performance. The second theme is Test as a Measure of the Student. The participants rejected the notion that high-stakes tests defined them as individuals or students. The third thematic element is Personal Relevancy of the Test. Five participants were critical of the test's connection to their current or future lives. The third element is Motivation. Only two students said they were not motivated to perform well on the tests, however, most participants said many of their peers were unmotivated by the assessments. The fourth element is Test Anxiety. Three participants said they were anxious when they took the tests. Two of these said they experienced memory blocks and physical symptoms. The fifth element is Human Connections and the Unknowable They. Most participants said

said they knew little about the reasons they were tested and about the people responsible for the testing process. They craved a connection with these unknown figures. The sixth thematic element is the Relationship to Silence. Four of the participants said they were uncomfortable with silence in general, and particularly with enforced silence during testing.

The researcher suggests further study of some important findings. These include investigating whether student test performance is affected by silence or by having a human connection with test designers.

A simulated panel discussion is used as a means to compare and contrast the findings of the study with legislative positions, the views of the researcher and the theories of John Dewey.

Timothy Pearson
School of Education
Colorado State University
Fort Collins, CO 80523
Spring: 2009

ACKNOWLEDGEMENTS

Writing a dissertation is a lengthy, difficult and predominately solitary experience. Without guidance and support from my committee, family and friends it is likely I would have never finished.

My committee members provided motivation and expert advice during the dissertation process. My advisor, Bill Timpson, provided me with critical and frequent feedback. Visits with Bill helped me to remain focused on my final goal. Jim Banning, my methodologist, possesses a remarkable ability to cut through the complexity inherent in qualitative research. He helped me to find my way when I was confused by many alternative routes. I also wish to thank him for encouraging me to push forward with the unorthodox structure I used in Chapter Five. Karen Spencer's constant encouragement helped me stay on course and her careful and thoughtful reading of my draft helped me craft a better final product. Cliff Harbor came late to my committee but his presence was critical in guiding me the final distance. He helped me to stay directed during the editing process and his suggestions transformed my fourth chapter into a readable work.

While my committee members helped me along on my academic journey, my family helped me persevere because of their emotional support. I was absent from many family conversations, social activities and shared recreation opportunities as I spent countless hours analyzing data and writing. I wish to take this opportunity to thank my

family for their love, support, and sacrifice during this long process. Their forbearance is more remarkable than my accomplishment.

To those listed above I would like to add my parents, sisters, brother and friends for their encouragement during this process. All of them nourished me and guided me through a process that would have been unspeakably lonely without their presence.

To Rebecca, Noah and Seth

CONTENTS

ABSTRACT OF DISSERTATION	iii
ACKNOWLEDGEMENTS	v
DEDICATION	vii
CONTENTS.....	viii
CHAPTER 1: INTRODUCTION	1
Background	1
Brief Overview.....	1
Rationale	2
Statement of the Problem.....	4
Research Question	5
Definition of Terms.....	5
Limitations and Delimitations.....	6
Researcher’s Perspective	7
CHAPTER 2: LITERATURE REVIEW	10
A Brief History of the Federal Role in Education	10
Early Federal Involvement in Education.	11
Lyndon Johnson and ESEA	11
A Nation at Risk.....	12
Texas Miracle.....	15
Writing of NCLB	17
Support for NCLB.....	21
Critical Views of NCLB	22
Reliance on a Single Test.....	22
Validity of Assessments.....	23
Effects on Minority and Low-Income Students.....	24
Strategies for Improving High-Stakes Testing Performance	24
The Student Testing Experience	26
Student Attitudes and Perceptions	26

The Role of Stress	27
Student Stereotyping	28
Student Cynicism	29
Testing and Purpose	30
Research Rationale	31
Progressive Educational Theory	32
Inclusive Leadership Theory	35
CHAPTER 3: RESEARH METHODOLOGY	38
Qualitative Research	38
Why Narrative?	39
Research Setting	41
Research Questions and Purposes	42
Selection of Participants	43
Data Analysis	46
Ethics	49
Trustworthiness	50
Personal Statement	51
CHAPTER 4: RESULTS	53
Introduction	53
Participant Overview	53
Narratives	54
Narrative #1: Thomas	55
Narrative #2: Dennis	61
Narrative #3: Michael	68
Narrative #4: Jackie	74
Narrative #5: Audrey	80
Narrative #6: Maria	87
Themes	92
School Loyalty and Test Compliance	92
Test as a Measure of the Student	94
Personal Relevancy of Test	95
Motivation	96
Test Anxiety	98
Human Connections and the Unknowable They	100
Relationship to Silence	101
Individual Participants and Thematic Elements	103
Thomas	104
Dennis	105

Michael	106
Jackie.....	107
Audrey.....	107
Maria	108
Summary	109
CHAPTER 5: DISCUSSION.....	111
Introduction.....	111
Discussion of Thematic Elements.....	111
School Loyalty and Test Compliance	112
Test as a Measure of the Student	113
Personal Relevancy of Test.....	115
Motivation.....	116
Test Anxiety.....	118
Human Connections and the Unknowable They	120
Relationship to Silence	122
Thematic Elements Summary	123
A Panel Discussion	124
Structure of the Responses.....	126
The Discussion.....	127
Citations for Preceding Panel discussions	152
REFERENCES	153
EPILOGUE	161

CHAPTER 1: INTRODUCTION

Background

In January 2002 President George W. Bush stood in a high school auditorium in Hamilton, Ohio with House members John Boehner R-OH, George Miller D-CA and Senator Edward Kennedy D-MA and signed into law the reauthorization of the Elementary and Secondary Education Act of 1965 ("President signs landmark education bill", 2002). In the seven years that have passed since that signing ceremony, this piece of legislation, known as No Child Left Behind (NCLB), has become part of the fabric of American public school culture. NCLB has radically changed the focus of public schooling in this country. States are now assessing students annually in math and reading.

In every state but Nebraska students are taking standardized tests to determine if their schools achieve the annual incremental goals set forth by the states under the direction of the U.S. Department of Education (Poetter, Wegwert, and Haerr, 2006. Steptoe, 2007). Because the tests are used to certify the performance of schools and districts, and because local school districts are required to publicly report their overall results, these assessments are frequently referred to as high-stakes tests. It is the students' response to these high-stakes tests that I am interested in investigating.

Brief Overview

I conducted a small-scale narrative inquiry using high school sophomores as participants. I questioned them regarding their testing experiences and about the test

preparation they received. Their narratives, and an analysis of these narratives, are at the center of my investigation.

Rationale

My purpose for conducting this investigation was to gain some understanding of the students' experience as they took high-stakes examinations required by the state under NCLB. Although many more may exist, I propose four reasons why I think the students' perspective needs to be captured. I will briefly introduce the four reasons here and in chapter two I will develop two of these reasons in more depth.

The most compelling reason to explore this topic is because of the lack of studies or articles related to the students' perspectives. My search of several databases yielded few works devoted to student attitudes regarding NCLB or to high stakes testing. Likewise, a similar search of the Dissertation Abstracts database revealed no theses or dissertations focused on these topics. Little research appears to have been conducted on this topic at the time I embarked upon this study in 2006. Whether this means researchers had not considered exploring students' responses to high stakes testing or whether they had not considered the exploration worthy of investigation could not be determined. I would like to assume the former. In either case, there appears to be a need for more information on this topic.

A second rationale for studying student perceptions of testing comes from the writings of John Dewey, considered by many to be the founder of the modern progressive education movement. Dewey was critical of imposing instruction upon students. Instead, he suggested educators should create rich and engaging experiences for students so they could learn themselves. He also believed students' prior experiences should not be

ignored when teachers engage with students. He felt the student's existing knowledge was frequently ignored by educators (Dewey, 1963).

Dewey, I believe, as an advocate of student-centered learning, would have been interested in understanding how students experience testing. As a proponent of student involvement in their learning process he would likely be interested in seeking student responses regarding this significant component of the schooling process.

A third reason for investigating student responses regarding high stakes testing also comes from a progressive tradition associated with a major discipline, in this instance, the field of leadership. A theorist from this discipline, Robert Greenleaf (1995), is the founder of a philosophy of leadership known as Servant Leadership. Greenleaf believes leaders have a moral obligation to lead in a way that displays they are willing to serve those they presume to lead. Since educators are certainly leaders, I believe this philosophical stance can be applied to them. One way educators could demonstrate their willingness to serve their students is through attempting to understand the experience they have when they are taking accountability tests.

The fourth, and final, reason I believe it is important for the voices of students to be heard is so the institutions involved in the testing process can benefit from their insights. There are several compelling reasons why teachers, administrators and school board members should hear what the students have to say about their testing experience. One reason is to simply to inform practice. I think schools should ask what they could do to make students feel more comfortable with the testing process. Another set of questions relates to test taking and test content. Has test preparation been so excessive that students are anxious or beyond caring? Are there areas where students need more preparation?

Schools can also learn from questions directed at specific groups of students. What are the reasons some students refuse to take the high-stakes tests? How do the attitudes of students of varying achievement levels, socio-economic level, age or ethnicity differ? These and other questions could help school districts make decisions regarding how they approach the exams with their students.

Statement of the Problem

Accountability testing dominates the focus of educators and administrators across the country. Many educational authors are critical of the effect this emphasis on testing has on students and on student learning. One of these authors is Kornhaber (2004). In a detailed and extensively supported article, she introduces a set of strategies she claims are used by educators to increase test scores. These strategies, which she refers to as “gaming,” share the characteristic of increasing test scores without increasing student learning. Three of these strategies are: Deliberately holding poor performing students back in the year before accountability tests are given, teaching test-taking strategies, and cheating on the part of teachers or administrators.

Each of these three strategies has a direct effect upon students and their schooling. Authors have suggested other negative effects of accountability testing on students. For example, some researchers (Posner, 2004, "President signs landmark education bill", 2002; Wolk, 1998) assert accountability testing has compromised the quality and depth of instruction. Another concern raised by researchers is that testing has negatively affected minority students (Davison *et al.*, 2004).

It is these concerns about the negative consequences of testing that stimulated my interest in investigating this issue through the eyes of students. If accountability testing

has negatively influenced teaching and learning, students are likely to have an understanding of these influences. If, on the other hand, students feel accountability influences have improved their educational experience, they should be able to communicate that as well. In either case I believe there is a great deal to be learned from the students regarding the influence of accountability testing upon their perspectives regarding learning.

Research Question

I began my study with a single research question: What do tenth grade students at Silverheels High School experience when they take the Colorado State Assessment (CSAP)?

Definition of Terms

High-stakes test – Standardized tests used for state or federal accountability requirements. The scores are used both to assess individual student ability and to rate the school according to NCLB requirements. In Colorado the tests are generally “high-stakes” for the schools but not necessarily for the students since their advancement and grades are not tied to the scores they receive on the exams.

No Child Left Behind Legislation – Reauthorization of the Elementary and Secondary Education Act (ESEA). Signed into law in January 2002 (Sunderman, Kim and Orfield, 2005).

Colorado State Assessment Program (CSAP) – State of Colorado criterion-based assessment tied to state educational standards and used to establish adequate yearly progress (AYP).

Adequate Yearly Progress (AYP) – Incremental annual academic performance targets for states, school districts and schools in reading, math and science towards the eventual goal of achieving full student proficiency by the 2013-2014 school year (Tracey, 2005).

Proficient – When applied to individual students in Colorado it means they have received a partially proficient, proficient, or advanced score on their CSAP examination. These proficiency levels are criterion referenced based upon skills and competencies designated by the Colorado Department of Education (n.d.).

Narrative Inquiry – “Research based on studying whole persons in context and in time through the narratives of their experience” (Josselson, Lieblich, and McAdams, 2003).

Limitations and Delimitations

I view some of the delimitations of my study not as barriers, but as welcome artifacts of my methodological choice. Just as financial investors trade risk for potential profit, I have traded some ability to generalize my findings to other settings for richness and complexity. Since my investigation is a narrative inquiry, I need to be cautious about making statements linking my findings to other educational settings, particularly schools with different demographics. That said, I hope the insights gained from my inquiry are deep and multi-faceted and have the capacity to inform others regarding the testing process.

Beyond the delimitations resulting from my chosen methodology, the school setting from which I drew my participant students presents barriers to generalizing to a larger population as well. My sample includes high school sophomores from a single school in Ute Creek, Colorado. This school has higher levels of economic and ethnic diversity than other high schools in the Ute Creek Unified School District, but lower

levels of low-income and minority students than may be found in other urban schools in Colorado, and significantly lower than in many high schools elsewhere in the nation. (School accountability reports, n.d.).

The age of my students may also have influenced my results. Some local educators and parents advised me that high school sophomores often adopt an apathetic attitude regarding the CSAP exam either because of their age, because of the number of CSAP exams they have already taken, or because of the finality of the assessment. I, however, chose sophomores for my investigation primarily because I expected them to be more articulate regarding their CSAP testing experience than younger students might be.

The limitations included protecting the anonymity of my participants and protections against undue influence of my participants during the interview process. I discuss the latter further in the methodology section under ethical considerations.

Researcher's Perspective

In March 1962 I sat in the Library of Denison Elementary School with fifty other fourth graders. Resting on the table in front of me were two freshly sharpened pencils and a pink eraser. At the front of the room the principal stood reading a prepared script. I looked out the window at the greening grass and the beige brick houses across the street as he droned on. When he finally stopped, a teacher handed me a test booklet and an answer sheet covered with circles and numbers. The principal began speaking again, explaining how to correctly fill in our names, our grade and the name of our teacher on the test forms. He spoke at length about the mysterious bubbles that ran down the page beneath each of letters in our name, about the importance of fully blackening each bubble and about the sin of making stray marks. When we were given the signal to begin testing,

an ominous silence fell across the room. The hands on the big clock above the library door moved at a glacial pace as I contemplated each of the questions and carefully blackened the circles that corresponded with the answers I thought were correct.

Part way into the testing session during the second day, I stopped caring. I quit reading the test items and began blackening the circles without any regard to whether they matched correct responses. My new challenge was to make the blackened circles appear to be randomly placed. It would not do, I knew, to have too many of the same letter blackened in the same column. Released from the drudgery of deciphering the unending questions, I felt light.

That experience as a test-taker is still with me. I have been a classroom teacher for almost 25 years and every one of those years I have been forced to play the same role my elementary school principal played so many years ago. As the first robins probe for worms and as the daffodils flutter in the March breeze, I hand out thick test booklets and read dreary scripts to my students. In the early years these were Iowa Tests of Basic Skills, standardized tests of shorter duration and much lower stakes than the CSAP exam I have administered since 1999.

I urge my students to do their best. I prepare them by teaching various test-taking strategies. I give them lemon drops and graham crackers. I lead them in stretch breaks half way through the test. Remembering my own experience, I avoid testing them for more than 55 minutes in a day. I do these things to make their task easier, but I am still haunted by the memory of that boy sitting in the library 45 years ago, chewing on his pencil, watching the clock that would not move and tormented by the fragrant spring air drifting in through the open windows.

I know all students do not share my testing experience. My two sons and most of my students take standardized tests very seriously. They carefully consider each question and fully blacken the correct circles or write well-crafted sentences in response to a short response questions. For these students the testing process is not arduous, but in every class I have ever taught sits at least one child for whom the clock does not move, someone who is tortured by the testing experience, someone who fills in the circles randomly – buying himself some relief from the testing pain. “Are you sure you are done?” I ask him when he finishes early. “Why don’t you go back and check your answers?” Invariably, within minutes, he has once again turned over his test on his desk and he is leafing through a Garfield book or taking apart a ballpoint pen.

As I conduct this study, the boy who thought he had mastered making random patterns on his answer sheets and the child who seeks refuge in a Garfield book are both with me. Although I am not a foe of standards, I am a reluctant participant in standardized testing. I know the discomfort a wiggly child feels as she tries to make sense of words she suspects were carefully crafted to trick her into supplying the wrong answer. I know the exhilaration that comes from abandoning the effort and letting chance determine the score.

CHAPTER 2: LITERATURE REVIEW

A Brief History of the Federal Role in Education

During the last 40 years tethers associated with accepting federal education funds have inseparably bound states and school districts to Washington. About 7% of the overall public education budget in the United States comes from the federal government in the form of Title I funds (Matthews, 2003). Although this percentage seems quite small, few school districts in this country can afford to forgo this income. It is this connection to federal funding that has allowed the federal government to impose mandates on local school districts through the No Child Left Behind (NCLB) legislation and previous versions of the Elementary and Secondary Education Act (ESEA).

The NCLB legislation is comprised of more than 1,100 pages and contains numerous provisions including early childhood reading programs, teacher and paraprofessional qualifications, and other lesser-known provisions (Robelen, 2002). All of these elements are important and have a substantial effect upon public education in this country, but for the purposes of this study I will focus upon the testing and accountability requirements that have driven high-stakes testing in most states.

Because of the substantial influence NCLB has on educational policy, and because the testing influences I wish to investigate are in a large part related to this legislation, I believe it is important to understand the events that led to the passage of this influential legislation.

Early Federal Involvement in Education.

The state departments of education, school districts and the federal government were not always this intimately linked. Historically, they had not been connected. Until the middle of the last century the U.S. government left the task of educating America's children to state and local governments. The role of the federal government had been chiefly that of a record keeper. It took an event from the cold, airless regions of outer space to convince the federal government to become involved in the nation's schools. President Eisenhower didn't panic when the Russians launched Sputnik, the first man-made satellite in 1957, but some of his advisors were uncomfortable with that symbol of Russian technological expertise circling overhead and many legislators shared their concern. Congress passed legislation Eisenhower later signed authorizing the spending of \$1 billion on educational programs designed to assure future American dominance in the areas of math and science. This marked the first time in U.S. history that large sums of federal money were spent on educational policy. The National Defense Education Act (NDEA) of 1958 opened the door for federal spending on education, and may have paved the way for more significant government policies in the decades to come ("The U.S. Response to Sputnik").

Lyndon Johnson and ESEA

Seven years after Sputnik, Lyndon Johnson identified a new crisis in America as he surveyed the poverty and lack of opportunity that characterized the underclass. To this complex problem he suggested a simple answer – better education (Tyack & Cuban, 1995). Johnson was the first president willing to use federal educational policy as a tool to equalize the educational landscape for disadvantaged children. This conviction, and the

passage of the ESEA of 1965, ushered in an era of increasing involvement in education under the leadership of presidents from both political parties.

As we examine the path that winds from Sputnik to No Child Left Behind, it is instructive to look at the focus of the first ESEA. The NDEA funds in 1958 were directed mainly at math and science education. ESEA funds in 1965 were predominately directed towards economically and culturally disadvantaged students. As part of Johnson's larger "War on Poverty," ESEA was to offer low-income students better opportunities through improved educational experiences. One billion dollars was spent on early childhood programs that impacted children of all races (Baylis-Heerschop).

Another significant development on the road to increased federal involvement in education was the establishment of a new post in the presidential cabinet, the Secretary of Education. Because Jimmy Carter came into office in 1977 with a promise to reduce federal government, he was not initially receptive to the idea of creating a new department. But Carter may have owed his nomination at the Democratic convention to a labor coalition that included the National Education Association, an organization enjoying rising influence in the political arena. When a hotly contested bill authorizing the formation of a department of education was placed upon his desk, Carter added his signature. Shirley Hufstelder, in her role as the newly-appointed secretary, toured the country spreading the word about the "good work classroom teachers do" (Stalings, 2002).

A Nation at Risk

The newly formed Department of Education appeared poised to be a casualty of budget cuts in the winter of 1981. Reagan shared a conservative disdain for federal

bureaucracy and had little interest in spending money on social agendas. One of his priorities upon coming to office was to terminate the new Department of Education. Given the task of dismantling the agency was Terrel Bell, Reagan's new Secretary of Education. Ironically, Bell set into motion a series of events that assured the continuance of federal involvement in education for more than two decades. Under Bell's leadership, federal involvement in education declined, but before Bell was able to pull the plug on the Department of Education, he released a report on the current status of America's schools. The document was effectively designated, *A Nation at Risk*. Bell's scathing report on the condition of the nation's schools shook the country. Suddenly Americans were hearing that their schools were mired in a "rising tide of mediocrity." (Stalings, 2002). Citizens were surprised to hear their schools being described as "failing." This report contained dreary statistics supporting the claim that the nation would for the first time produce a generation of learners whose abilities would not match those of their parents (Tyack & Cuban, 1995).

The furor created by this report may have caused Reagan to reconsider eliminating the department. The Secretary's position gave him an opportunity to push forward one of the pet agendas of the conservative movement. *A Nation at Risk*, with its pessimistic view of the national educational system, provided a perfect backdrop for the Republican president to unveil the idea of school choice (Fowler, 2004). Conservatives hoped the public would come to share their notion that healthy competition would lead to superior schools.

"A Nation at Risk" assured the continuance of the Department of Education, but it also was the catalyst for the national standards movement that eventually led to the No

Child Left Behind legislation. President George H.W. Bush in 1989 convened the first National Education Summit in Charlottesville, Virginia. Bush, and the 50 governors attending the conference called for the formation of national educational standards. They also set forth six goals to be achieved by the year 2000. These goals, known as the *Goals 2000*, proposed several lofty aims, among them the attainment of 100 percent adult literacy, 90 percent high school graduation, and reading readiness for all children starting elementary school (Young).

Although Bush lost his bid to return to the White House in 1993 his educational efforts continued. Bill Clinton had participated in the National Education Summit as the governor of Arkansas. As president, he stepped up his involvement in the national standards movement by signing the Improving America's Schools Act of 1994 (IASA). This was a reauthorization and revision of the 1965 ESEA. Under this legislation states were required to implement challenging content standards in mathematics and language arts; come up with performance standards; develop assessments tied to the standards; use the same standards to measure Title I students that were being used for measuring other students; and establish improvement benchmarks, known as adequate yearly progress (AYP) (Young, n.d.).

Late in Clinton's second term, nine months before the inconclusive election that eventually led to George W. Bush becoming president, Congress was again looking to reauthorize the 1965 ESEA legislation. Bill Gooding, a Pennsylvania Republican, former teacher and school administrator, was the Education and Workforce Committee chair. Having declared that he was not seeking a 15th term in the House, his wish was to leave Congress with the latest reauthorization of the ESEA as his legacy. James Jeffords was

the Republican head of the Senate Education Committee. Jeffords had frequently irked conservatives in his party by not going along with their educational agenda items such as vouchers (Sack, 2000). Serving with him on this committee was the experienced and powerful Edward Kennedy.

Some of the Republicans were pushing for a major change in the ESEA while others were willing to compromise. With the presidential campaign in full swing, the stage was not set for a harmonious bipartisan effort, as had been the norm in previous reauthorizations of ESEA. Congressional representatives were primarily split along party lines on issues surrounding reauthorization. Republicans were using the term “flexibility” to describe their view of how states should be permitted to spend Title I funds. By this they meant, although it wasn’t always clearly stated, they wanted states to be able to spend federal funds on vouchers, private tutoring, and a provision to pay to transport children from failing schools. The Republicans were also looking for accountability measures and were pushing for block grants. The Democrats, for their part, wanted increased spending, class-size reductions, and school building renovation to be part of the authorization. The “Third Way,” a compromise position fostered by some conservative Democrats, was not moving forward during the session. The differences proved to be insurmountable in the 2000 election year, and ESEA ended up being pushed aside so more pressing issues could be brought forward. ESEA would have to wait for the next Congress (Robelen, 2000).

Texas Miracle

While Clinton was pushing educational standards at the national level, the nation’s governors were at various levels of involvement in efforts to introduce

educational standards within their home states. George W. Bush capitalized on the efforts of two previous Texas governors to move his state beyond the requirements of IASA. The Texas model included a statewide exam, known as the Texas Assessment of Academic Skills (TAAS), an educational report card that displayed the performance of individual schools and districts, and sanctions for poor performing schools. These sanctions were relatively mild. They required improvement plans for schools that posted poor scores and offered the parents of students in low-performing schools the option to transfer their children to schools with higher test scores ("Introduction to the Texas accountability system," 2002).

Other educational changes were attracting attention in Texas at same time statewide reforms were being implemented. Rodney Paige, a former college dean, was the no-nonsense superintendent of the enormous Houston School District. Under Paige's tough policies, principals who couldn't deliver test score increases and near-zero student dropout rates were fired. Principals who did deliver got yearly bonuses as large as \$5,000. The test scores did go up, and many schools in Houston reported no student dropouts. Despite skepticism on the part of many teachers and academics, the successes in the Houston schools, and gains throughout the state, became known as the Texas Miracle (Becker, 2003).

George W. Bush was running for president in 2000 the same year, according to Goals 2000, that was supposed to see universal literacy and high school graduation rates in excess of 90%. George W. Bush wanted to change the public's image of the Republican Party. As part of his push for "compassionate conservatism" he was advancing his educational agenda on the campaign trail (Magnet, 1999). Unfortunately

for Bush the Rand Corporation released a report critical of the Texas schools shortly before the election. “It’s not a miracle,” asserted Stephen Klein (2000), chief Rand investigator, assessing the claims of impressive gains in the Houston schools, “We think these scores are misleading and biased because they’re inflated.” The Rand study criticized schools for drilling students extensively to prepare them for the state exams. It also pointed to dropout rates for minority students estimated at 50%. Bush, just two months from the November elections, stood by his record and reasserted himself as the educational candidate.

Writing of NCLB

Despite the uncertainty of the presidential election and a lack of a clear mandate, Bush was uniquely poised to push his educational agenda. He had the support of the nation’s governors, key Democrats, his own party, and the public. Resistance to state standards had waned, and many states were already accustomed to the idea of high stakes testing. The little resistance to testing and national standards that remained hardly posed a significant obstacle. Most of the wrangling that occurred in the Congress concerned funding issues rather than policy issues.

The 2000 election introduced new players to the ESEA reauthorization process. George W. Bush, the “education president” had the leading role. He brought a new plan for educational reform. Jeffords survived his reelection bid and was back as head of the Senate Education Committee. Kennedy’s return was a given. The Republicans were dominant in the House, but the Senate was split between the parties. Bipartisan support would be necessary to get ESEA to the President’s desk.

Bush's plan resembled what the Republicans had been pushing in the previous Congress. Included in the proposal were provisions for annual testing, tough accountability, flexible spending, block grants and vouchers. The Democrats were still looking for class-size reduction, school renovation and increased spending. In May, the ranking Democrat on the House Education and Workforce Committee, George Miller, acknowledged that the bill being crafted in his committee was a compromise. The Republicans in the Senate also conceded that they had taken a compromise position as they dropped vouchers from their bill. In the Senate version of the bill Democrats gave ground on class-size reduction and lost their school renovation component. Funding levels remained as the biggest sticking point in moving either the House or Senate toward a final vote (Robelen, 2001a).

By late May the House had passed their version of the education bill. Preceding its passage, an unusual alliance attempted to remove the testing provisions from the language of the bill. Representative Barney Frank (D-MA) joined Representative Peter Hoekstra (R-MI) to introduce an amendment that would have removed the annual testing provisions. The conservative Hoekstra, referring to news stories detailing mistakes in the design and scoring of standardized tests said, "Testing is not ready for prime time." The amendment failed and testing remained in the House version (Fine, 2001).

Jim Jeffords of Vermont upset the balance of the Senate on June 5 when he switched from the GOP to independent status. The Democrats now held a single seat majority. Jeffords lost his chair of the Education Committee to Edward Kennedy (D-MA), but he said he had little influence anyway because his views had differed so markedly from those of his party. Despite the change in leadership, and disagreements

about spending levels and vouchers, the Senate passed the ESEA reauthorization bill on June 14. The Senate bill included mandatory testing provisions, accountability requirements for states, mandatory spending requirements for children with disabilities, full funding for programs to educate poor and disadvantaged children, and an expansion of bilingual programs ("Senate passes education bill," 2001).

Since the Senate and the House had each passed versions of the ESEA legislation, a compromise had to be crafted in the conference committee. Significant differences between the Senate and House versions made the task of the committee a difficult one. Differences included levels of funding for special education, funding for overall education, and how to define school progress (Olson & Robelen, 2001).

With the bill in the hands of conference committee, the testing opponents decided to speak out. Little had been heard from this camp before. Lobbyists from various public interest groups and organizations thought the time was ripe for affecting a change in the provisions regarding high stakes testing. Bob Chase, president of the National Education Association, claimed in his address at the group's annual meeting "The soul of education is being jettisoned in the name of testing." (Blair, 2001).

As Congress moved towards its summer recess, Bush cautioned lawmakers against setting "impossible expectations." The Congressional Research Service issued a report in July indicating Maryland, North Carolina, and Texas would have had few schools meet the adequate yearly progress requirements in either the Senate or House versions of the education bill if they had been applied to their 1998-2000 results. Roy Romer, superintendent of the Los Angeles schools and former Colorado governor, expressed his concern that the bill was too "prescriptive." Senator Wellstone (D-MN)

pushed for the use of multiple measures of performance (Blair, 2001). In the end, their dissent went unheeded.

The terrorist attack of September 11, and the ensuing anthrax scare, prevented the conference committee from meeting until November. Disagreements about access for Boy Scouts to school buildings and funding hate crime programs were eventually worked out. The committee agreed to use the National Assessment of Education Progress as a benchmark to gauge state performance. The test was to be given every two years to a small sample of fourth grade students. Disagreements over funding for special education persisted (Robelen, 2001b).

By December 2001, the committee finally was able to agree on a compromise bill. The Democrats wanted more funding, but settled for a 4 billion increase in educational spending. The Republicans prevailed, preventing a move of special education funds from discretionary spending to mandatory spending. The Republicans also got their provision allowing private tutoring. As the two legislative bodies prepared to vote on the final bill, the American Association of School Administrators and the National School Boards Association opposed its passage on the grounds that it lacked mandatory special education funding. The National Conference of State Legislatures expressed its opposition to bill because of “unfunded mandates” (Robelen, 2001d).

The House passed the compromise bill with a vote of 381-41 on December 13. The Senate followed suit on December 18 and the bill went to President Bush to sign. Bush, in a dramatic public ceremony attended by Representative George Miller (D-CA), Senator Edward Kennedy (D-MA), Representative John Boehner (R-OH), Secretary of Education Rodney Paige and some students, signed the reauthorization of ESEA into law

on January 8, 2002 in a high school auditorium in Hamilton, Ohio, Boehner's home town ("President signs landmark education bill", 2002).

The harmony displayed on that day in Ohio soon dissipated. On March 18, 2002 George Miller, ranking Democrat on the House Education and Workforce Committee and Edward Kennedy chair of the Senate Education Committee, sent a letter to President Bush asking him to honor his commitment to fund education at the level designated in the Act. Miller and Kennedy complained the President's budget proposal cut funding for NCLB by \$90 million, while at the same time it proposed tax cuts for the top one percent of taxpayers in an amount totaling 50 times that sum. Kennedy and Miller (2002) called on Bush to back funding increases to assure the success of this historic school reform. The bipartisanship that had characterized the drafting of NCLB began to unravel. The funding issues that had been characteristic of the legislative negotiations reemerged as the chief difference between their stands on education.

Support for NCLB

Support for NCLB among conservatives has historic roots. Ross Perot pushed for standards and a statewide test in Texas in 1984. The Texas system later became a model for the national system proposed under NCLB (Kornhaber, 2004). Other elements of the legislation are similar to those found in a system developed under the direction of Tom Ridge when he was governor of Pennsylvania. His plan provided incentives and penalties and provided for the takeover of "failing" schools (Pipho, 1999).

The views of many political conservatives regarding accountability are evidenced in Samuel Carter's book, *No Excuses* (2001). Published by the Heritage Foundation, this report showcases 21 high poverty, high performing schools from different regions in the

country. The book begins with a blistering condemnation of the status of America's schools by way of an introduction by Adam Meyerson, Vice President, Educational Affairs at the Heritage Foundation. Meyerson begins with a statement reminiscent of "*A Nation at Risk*" when he says, "America's public schools have utterly failed." He goes on to cite statistics to support his assertion. He hands the baton off to Carter who continues the case. Carter directs his blame squarely at the educators in this country who have coddled low-performing students by expecting too little from them. He makes his case for higher standards, more accountability and better programs. Carter expresses little patience for the excuses offered by liberal educators for the low performance of poor and minority children. He points to schools featured in the book as proof that poor and minority children can excel given high expectations.

Support for NCLB is not restricted solely to conservatives, some groups that advocate for minority and special education students positions are also supportive. Their support is tied to a provision in NCLB requiring that all sub-groups make progress, including minority, low-income and special education students (Robelen, 2002).

Critical Views of NCLB

Reliance on a Single Test

Although many states recognize the need for multiple measures of student achievement, the accountability and assessment portion of NCLB relies upon single annual examinations. Some states have dual accountability systems; one they believe represents assessment best practices and another that satisfies the requirements of NCLB (Sunderman, Kim and Orfield, 2005). In the Standards for Educational and Psychological Testing (Gulek, 2002), one of the most important testing standards

outlined is one that states “Any decision about a student's continued education, such as retention, tracking, or graduation, should not be based on the results of a single test, but should include other relevant and valid information.” Neill (1996) calls for assessment based upon a “complex of classroom-based assessments revolving around observation, documentation and evaluation” (p. 5). His vision of assessment includes gaining a multi-faceted view of the learner using many forms of assessment. He suggests the chief purpose of these multiple measures should be to guide effective instruction. By using evidence of learning collected over time, and by using different assessment means, he thinks it is less likely that a high stakes decision will be made in error. Another advantage to using multiple assessments, according to Gulek, is that a more complex reconstruction of the strengths and deficits of the learning experience can be achieved; making it more likely changes in pedagogy can improve instruction.

Validity of Assessments

Hoover (2003), a Youngstown University researcher, analyzed data from 593 Ohio districts that administered the Ohio Proficiency Test (OPT). Using available economic data, he developed a multiple variable indicator of socioeconomic status (SES) he calls the Presage Factor. Hoover compared Presage Factors to OPT scores. He found a very strong correlation (0.8) between test scores and economic advantage as measured by the Presage Factor. This strong association led Hoover to question the validity of the test. He questions whether the OPT is measuring student learning or economic status of the neighborhoods where the students live. A correlation of this strength makes it difficult to ignore this relationship.

Effects on Minority and Low-Income Students

Texas, like Ohio, has a high stakes assessment tied to state standards that has been in use for several years. The Martinezes (2002) found an increase in attrition rates among Texas Hispanic high school students. Attrition rates increased from 45% in 1986 to 52% in 2001. Since the 1986 date preceded accountability testing, they suspect there may be a relationship between accountability testing and minority dropout rates. Another trend they noted was the retention rate for Hispanic and Black ninth graders. This number was 30%, much higher than national means. The Martinezes speculate these high percentages reflect a deliberate policy of holding back ninth graders so they will not have to take the required TAAS exam as tenth graders. The Martinezes claim grade retention is the single greatest predictor of a student dropping out of school. In this way fear of sanctions from the high-stakes testing and the associated high retention rates may have contributed to pushing out the very students whom the policies are targeting.

Strategies for Improving High-Stakes Testing Performance

Under pressure from federal and state governments to produce favorable test scores, many schools have employed strategies to boost students' scores. Kornhaber (2004), in a detailed and extensively supported article, introduces a set of actions used by educators to increase test scores. She calls practices intended to increase test scores without increasing student learning – gaming strategies. A description of three of these practices follows.

The first activity identified by Kornhaber (2004) is the practice of changing the nature of the pool from which the test-takers are drawn. Some safeguards were built into NCLB to discourage this activity by requiring high percentages of students to take the

test in each school. Another safeguard is the requirement that minority, special needs, and low-income groups be viewed separately rather than averaged with the majority population. Not prevented by the legislation according to Kornhaber, is the practice of retaining students likely to perform poorly in the grade before they are tested. This is the practice the Martinezes (2002) believe they identified in Texas.

A second gaming strategy she identifies is the practice of teaching test-taking skills to students. These skills include teaching students how to fill in the bubbles and providing them with strategies for eliminating untenable responses. Instructing students in test-taking skills raises the scores, according to Kornhaber, without raising student knowledge. This instruction creates savvy test-takers but may subtract from the time that students are instructed in other content.

The first two gaming strategies may be ethically suspect, but the third is clearly fraudulent. Cheating is the ultimate act of gaming, and although its direct effect upon student learning is more difficult to determine, it certainly sends an unfortunate message to students. Cheating may include underreporting dropouts as was revealed to have occurred in the “Texas Miracle.” According to Becker (2003) a state audit in Texas revealed that nearly all of the 16 schools audited had underreported their number of dropouts. By allowing students likely to reduce test scores to leave without reporting them, the test scores increased and the reported dropout rates did not reflect an increase.

Examples of cheating were also reported in the school district in which I am employed. Two Poudre School District elementary principals were investigated for violating state accountability rules. According to a report in the Fort Collins Coloradoan (Lingle, 2005), one of the principals gave students additional time to complete the state

assessment; the other revealed test items to teachers and paraprofessionals and encouraged them to ensure the students knew the material prior to administering the test.

The Student Testing Experience

If a student taking a high-stakes test resembles a pebble tossed into a pond, then I have found little information in my review of the literature about the initial splash created by that pebble, but much about the distant ripples it produced. As one moves from the central event, the circumference of the ripples increases. The challenge is in limiting the scope of the investigation. As I look closer to the center, where the rings are tighter; there is less information available. In this section I will first present the findings from a single study that looks at the splash of the pebble – the student experience. Next, I will present data that describes phenomena closely related to the experience, but a ripple or two removed from the central event.

Student Attitudes and Perceptions

Mulvenon, Connors, and Lenares (2001) conducted the sole study I found in the literature that focuses on the student testing experience. This study investigated the relationship between student attitudes and perceptions and test performance. The researchers surveyed 283 Arkansas fourth graders and compared the survey data with the students' results on two state-mandated tests. The first test was the norm-referenced Stanford Achievement Test; the second assessment was the criterion-referenced Fourth Grade Benchmark Examination. They found participating students generally experienced little stress related to the exams. They also found no evidence anxiety or pressure exhibited by teachers or parents impacted student test performance. Most of the surveyed

students had positive attitudes regarding testing. Additionally, researchers found students who held favorable attitudes regarding testing had higher test scores.

These are controversial findings. The authors admit they run counter to the beliefs of many educators. I think two factors must be kept in mind when evaluating their findings. First, the students surveyed were fourth graders, and as such were relatively new to a high-stakes testing environment. Would older students have been as sanguine about testing? Second, the student participants who formed the sample for this study differ from the non-participants in one important regard - they returned the survey. All of the fourth grade students in the Arkansas district were given the survey. The researchers had a survey return rate of about 50%. I think it is possible those students who failed to return the survey might hold less favorable attitudes regarding the exams. These two factors should certainly be considered when evaluating the results of this study.

The Role of Stress

Other studies have found contrary findings regarding the role of student stress and testing. One of these studies ("Wake-up call over teen stress") looked at high school students in Northern Ireland. This study, conducted by the University of Dublin and Queen's College, investigated teen stressors. They found 75% of the respondents were stressed by schoolwork at least some of the time; with 25% saying they were "often" or "always" stressed by the schoolwork. The researchers link this stress to school examinations. They were surprised to find students were more likely to be stressed by testing than by bullying.

Another study more closely examined student stress related to testing. Beilock and Carr (2004) investigated a phenomena noted by other researchers who had found

highly competent undergraduate math students performed more poorly in high anxiety situations than low-anxiety situations. Beilock and Carr's study was designed to test the hypothesis that anxiety interferes with working memory. They found participants who scored highest on measures of working memory were the most likely to suffer performance declines in high-pressure circumstances. The researchers point to the irony that those who have the highest ability levels were the most likely to see performance declines in high-stakes testing situations. Students with lower measured levels of working memory were less likely to show performance declines in the same situations.

If true, this calls into question the validity of many high-stakes assessments, including the Scholastic Assessment Test (SAT), and state examinations tied to NCLB requirements. If the ability of our brightest and most capable students is not measured accurately by high-stakes exams, do they measure what they purport to measure? These findings provide another argument against the viability of high-stakes exams, but they also supply important background information regarding the student testing experience.

Student Stereotyping

Like stress, student perceptions regarding stereotypes may play a strong role in their performance on standardized tests. Good, Aronson, and Inzlicht (2003) designed interventions to overcome racial and gender stereotypes among Texas seventh grade students. College mentors were taught to work with the students, providing them with one of two persistent messages. The first message was that intelligence is malleable, something that is not static or set. These students were encouraged to think they could improve their intelligence through effort. The second message conveyed by the mentors

was that some students initially struggle, but that they could bounce back from their difficulties. Both messages shared the same theme of persistence despite difficulties.

Both intervention groups performed significantly higher on standardized tests than the control group. This led the researchers to conclude that mentors, providing strong counter messages, can overcome racial and gender stereotypes. One of the targeted stereotypes was that students of color do not perform well on tests because they are not as capable. Another was that boys are better than girls at mathematics. The significance of this study is that it suggests student test performance is influenced by student-held stereotypes and that these stereotypes can be overcome by persistent counter messages.

Student Cynicism

Some of the literature critical of high-stakes testing decries the use of deceptive testing practices. I presented some of these activities Kornhaber (2004) identified as gaming strategies earlier in this chapter. If some educators engage in deceptive, or disingenuous practices, do students notice?

Two Yale researchers, Mills and Keil (2005) may provide some insight into this question. Most of us tend to think of children as naïve and trusting, and according to Mills and Keil many social scientists share that view. They conducted a study that contradicts this notion. The researchers found children, particularly children older than second grade, attributed behaviors to self-interest when they analyzed the actions of characters in short stories. The researchers concluded children might be at least as cynical regarding the intentions of others as adults are. The children in the study were as willing to attribute the actions of the characters in the short stories to insincerity and self-interest, as they were to attribute their actions to a simple mistake.

If these results accurately reflect potential cynicism in children, then students are likely, even at a young age, to be able to evaluate the motivations of teachers and administrators regarding the administering of high stakes testing.

Testing and Purpose

Seitsinger, Felner, Brand, Shim and Dumas, (2005) found parent involvement in students' education influenced their academic achievement. Those who received homework help from their families; those who participated in educational enrichment activities, and those who had families who conveyed a sense for the importance of education to students were found to perform better on measures of accountability.

Coppola (2005) wondered just what message students internalized from their families or other sources. Did they equate elevated test scores with being well educated? Had high stakes tests become a proxy for other educational goals such as achieving literacy, attaining knowledge, or gaining the skills needed to process knowledge? His interviews of students indicated they had indeed internalized testing as the purpose of schooling, but another finding was of more importance to him. The students almost universally cited the purpose of education was to provide for the means of fulfilling the American Dream. Students from many cultural backgrounds stated that the purpose of education was to provide them with the tools they needed to enable them to succeed in the workplace.

I am less surprised to hear this message than Coppola, perhaps because I have heard this frequently from my students. I think the powerful message here is not that schooling is perceived to be the means for securing a job, but that it is thought to be the

link between testing and being credentialed to obtain the job. If high-stakes testing is linked, at least in students' minds, to their economic future, is it worthy of that role?

Some would disagree that economic gain is what students want from education. Harvard researcher Unger (2003) conducted video interviews with students. The message he heard them convey was that they had a need to be engaged in their learning through discussions, projects, and experiences. They also said they wanted to participate in learning experiences that were relevant to their lives. It is not clear if by this they mean to their future economic welfare or to something more immediate.

Unger also discusses students' desire for involvement. He claims students crave more voice. He says that they want to be able to express their opinions, ideas and beliefs. He links business theory to educational settings, particularly client-centered approaches. In education the client is often considered to be the parents of students. Unger thinks the true clients are the students. Beyond stating that they desire more voice he does not attribute any specific motivation to students.

These theories regarding student purpose differ in their basic premises. One set of theories posits students desire to achieve success because of familial expectation the other suggests students want to learn so they can express themselves and use newfound skills to enrich their lives. Perhaps both sets of theorists are correct. I believe it is likely students find reasons to learn from multiple sources.

Research Rationale

I briefly explored four reasons for conducting my study in the introduction. The first reason I gave was that there are few studies and little literature related to this topic. The second reason was tied to the beliefs of progressive educators such as Dewey and

Palmer. A third reason looks outside of education to leadership theorists for support. My fourth rationale is to inform those in education about the student testing experience. In the following section I will further develop arguments for the second and third rationale.

Progressive Educational Theory

When I introduced my argument in support of including student voices in the NCLB discourse I evoked progressive educators for support. One of those educators, John Dewey, sought to transform an educational system that in many ways, as Tyack and Cuban have observed (1995), resembles our current public school system. Dewey viewed traditional education as an imposition of ideas from above and from the outside upon students. “It imposes adult standards, subject matter, and methods upon those who are only growing slowly towards maturity” said Dewey (1963) over sixty years ago. Dewey criticized a system that little considered the experiences of the student, one that imposed knowledge from above. He made a case for learning based upon experience and student interest. It is a mistake, according to Dewey, to provide students with knowledge deemed useful in the present without considering what students may need in the future. According to Dewey no educator is capable of predicting what the future needs of students may be. He wished to see students equipped with the tools needed to acquire knowledge rather than knowledge itself.

Dewey would likely be surprised by the nature of public education today. Despite the passage of more than a century since he published his earliest works on educational reform, few of the progressive elements he championed in public education have found a lasting place in the nation’s schools. *A Nation at Risk* and the standards movement have

shifted the emphasis in schools away from student-centered education envisioned by Dewey towards a model focused upon accountability (Stalings, 2002).

A modern progressive educator, Palmer, (1998) also considers the experience of students critical. An unpleasant experience students frequently encounter in the classroom, according to Palmer, is fear. “Students too are afraid of failing, of not understanding, of being drawn into issues they would rather avoid, of having their ignorance exposed or their prejudices challenged, of looking foolish in front of their peers” (p. 37). Palmer wishes we would turn the energies directed to school reform towards trying to understand the fear of students and the similar fear experienced by teachers. Palmer believes the indifference and hostility teachers sometimes encounter in their teaching is a response to fear students feel. When this fear is met with a teacher’s fear of being ignored, or of failing to reach every student, an unproductive cycle results that is antithetical to the learning process.

Palmer echoes Dewey on the topic of teacher-directed learning. He believes a conventional learning environment is centered on the premise the instructor will deliver conclusions to students. It assumes, according to Palmer, the teacher has all of the answers and the students have little or none. Palmer, though, is equally critical of a system that puts the student at the center of the learning environment. In this way he differs from Dewey. Rather than a student-centered classroom as Dewey envisioned, Palmer sees the ideal classroom as subject-centered. While at first this may sound like the very thing the architects of educational reform suggest, his vision is quite different. Palmer wants to see the best elements of student-centered and teacher-centered education merged with the subject at the center. This combination of academic rigor and strong

student and teacher involvement leads to a “community of truth” contends Palmer. The teacher’s task in the subject-centered classroom is to give the subject, which Palmer calls “the great thing” (p. 107), an independent voice. By this he means the subject subsumes the voices of the teacher and the students and becomes central to the focus of the learning process.

These two progressive educators have plenty to say about the nature of traditional education, but what would they say about testing, or more specifically, what would they say about the student’s role in the dialogue concerning testing? *A Nation at Risk*, the document that initiated the standards movement (Stalings, 2002) was released decades after the death of John Dewey. Nevertheless, I think it is safe to say that if Dewey were alive today he would reject high-stakes testing out-of-hand. If students were truly at the center of their own learning, if they were really directing the content and direction of their learning, a standardized test would not be an appropriate instrument to measure their learning. Standardized testing goes against what progressive education represents. Continuing with that line of thinking, Dewey, confronted with high-stakes testing, would likely maintain students’ input regarding their experiences with the testing should be solicited.

Palmer’s discussions can also be related to the dialogue concerning high stakes testing. I think the fear Palmer discusses is inextricably linked with high-stakes testing. The student is afraid of failing, the teacher is afraid he or she will not adequately prepare the students, and the administrators are fearful their schools will not get a high rating as a result of testing. Parents, school board members and community members also experience different levels of apprehension regarding the process.

Palmer's "community of truth" may be related to high-stakes testing as well. Where Dewey would likely consider how testing effected the learner, Palmer would consider how the test impacts the balance sought in the community of truth. That is to say he would likely concern himself with the way the test affected the central status of the subject. If the test results became primary, or if the anxiety of the learner and the instructor become a distraction, then the subject is no longer central. The community of truth becomes a community of fear, a community fixated on accountability, rather than on truth.

Although it is unlikely soliciting student opinions regarding high-stakes testing in itself will diminish the fear associated with the testing process, or that it will restore the primacy of content learning to the teaching process, perhaps an understanding of the dynamics of testing from the student perspective would improve the process.

Inclusive Leadership Theory

The third rationale I mentioned in my introduction for seeking student input regarding high-stakes testing comes from the philosophies articulated by those who advocate inclusive leadership. In a manner similar to progressive education, inclusive leadership models shift the focus away from the leader and towards the group. Lao Tzu (1995) in the Tao Te Ching said, "Remember that you are facilitating another person's process. It is not your process. Do not intrude. Do not control. Do not force your own needs and insights into the foreground" (p. 71). Greenleaf (1995), the originator of the leadership philosophy known as Servant Leadership, discusses similar leadership qualities. He speaks eloquently of a new form of leadership where coercion is not wielded as a tool of authority. "A new moral principle is emerging which holds that the

only authority deserving one's allegiance is that which is freely and knowingly granted by the led to the leader in response to, and in proportion to, the clearly evident servant stature of the leader"(p. 20). According to Greenleaf, only those who have proven to have the best interest of those they presume to lead will be recognized as leaders. Leaders should have proven their desire to serve first, before they choose, or are chosen, to lead.

Greenleaf's principles are evidenced in one of the processes examined by Wixson, Dutro and Athan (2003). They contrast the process adopted by two states in their creation of educational standards for public elementary and secondary schools. In Texas the process was dominated by office of the Governor and a few hired consultants. Educators, the public, and academics were excluded from the process. In Wisconsin, a collaborative model was employed where multiple and diverse voices were involved in the process. The authors label the two models respectively, the competitive process and the deliberative process. These two models reflect contrasting attitudes about leadership in general, and specifically leadership as it relates to education.

A competitive model similar to that used to create standards in Texas was used at the national level to produce the NCLB legislation. The White House provided a draft to Congress in the spring of 2001 (Robelen, 2001c) that was a template for reauthorizing the Elementary and Secondary Education Act (ESEA). This draft bore the philosophies that had been recently expressed by the Heritage Foundation (Carter, 2001), conservative educational scholars such as Diane Radvich (Stossel, 2000) and legislation already in place in Texas (Wixson et al., 2003). The process of reauthorizing ESEA, despite its bipartisan nature, included little input from educators or academics. Lobbyists representing educators had directed their efforts almost exclusively towards blocking the

inclusion of vouchers in the bill (Blair, 2001). In this competitive, rather than deliberative process, the fight over vouchers and funding levels distracted the legislators from debating the issues of accountability and educational improvement.

I would characterize the process that produced NCLB as not only competitive, but also paternal. It certainly bore little resemblance to the leadership philosophies articulated by advocates of participant leadership such as Robert Greenleaf (1995). The legislators and the president believed the best solution to boosting the achievement levels of low achieving schools was an accountability model including both reward and punitive elements. To this end a majority of both parties were united. The process of writing the legislation, although characterized by compromise, was not collaborative because the voices of researchers and educators were not included.

It is too late to change the procedure that was used to formulate NCLB. It is too late to include input from the various stakeholder groups. It is not too late, however, to solicit the opinions of these stakeholders and include it in the literature. This will not make the process inclusive, but it will at least give voice to players not previously heard. NCLB is not monolithic; pressure from some groups has already led the Department of Education to make changes at the national level (Robelen, 2004). Changes can be effected at the local and state level as well. This is not participant leadership, but it may lead an autocratic hierarchy to recognize the voices of some of the participants, including the students.

CHAPTER 3: RESEARCH METHODOLOGY

Since I took my first undergraduate psychology class over 30 years ago I have been uncomfortable with attempts by social scientists to explain the complexities of human behavior using scientifically based experimental designs. I agree with Gergen and Davis (2003), when they say of quantitative investigations, “The average of everything may not be a description of anything” (p. 242). When I first heard about the qualitative paradigm for investigating human behavior, I knew this was a means of investigation that I could embrace. Here was a method of investigating human behavior that recognized its depth and intricacy, a means that did not rely on connecting human activity to numerical values. This is not to say that I discount quantitative investigations out-of-hand, it just means that I am more comfortable examining human behavior using the methods of a qualitative investigator.

Qualitative Research

Creswell (1998) cites eight reasons why a researcher should choose to use a qualitative approach for an investigation. 1. Use a qualitative approach because of the nature of the research question. A qualitative approach should be used when it is important to understand how or what is happening rather than why something occurs. 2. Choose a qualitative study because the topic needs to be explored. By this he means the variables have not been identified and theories have yet to be constructed. 3. Use a qualitative approach when a detailed view of a topic is needed. 4. Choose a qualitative study when one wishes to study people in a natural setting. 5. Select a qualitative

approach because one has an interest in writing in a literary style. 6. Select a qualitative study because sufficient time and resources are available for data collection and analysis. 7. Opt for a qualitative study because the targeted audience is receptive to this type of approach. 8. Choose a qualitative method when the researcher is willing to accept the role as an active learner rather than as an “expert.”

In my opinion Creswell’s list provides a good litmus test for an inexperienced researcher who is embarking on a large-scale research project. Although I can apply all of the conditions to my study, I think Creswell’s first two reasons for choosing to use a narrative approach are particularly appropriate. The questions regarding *how* and *what* are central to my research question. I wish to understand and describe the experiences of students rather than explain the reasons for their responses. Little information is available related to my topic, meaning little is known about the phenomena of student responses to high stakes tests. For this reason, the topic needs to be explored as Creswell suggests in his second criteria.

Why Narrative?

I have chosen to conduct a narrative inquiry because it suits my investigation and because it suits me. Connelly and Clandinin (1990) say a narrative study “... is the study of the ways humans experience the world” (p. 2). According to them the researcher can come to understand the holistic nature of human experience through narrative inquiry. Josselson, Lieblich and Mc Adams (2003) expound on the role of narrative in revealing experience, “The qualitative/narrative researcher eschews methodolatry in favor of doing what is necessary to capture the lived experience of people in terms of their own meaning-making” (p. 260). This is precisely what I wished to do – to explore the

students' testing experience in all of its complexity using their narratives as an investigative tool.

Another reason to use narrative is because I am comfortable with the use of stories. My study could be just as easily conducted as a phenomenology or as a case study, but I am strongly drawn to the centrality of story in the narrative approach. In my 25-year practice as a teacher I have often been struck by the power stories hold for students. I have exploited that power, using stories in my instruction. I have taught the difference between area and volume using allegorical tales, I have written instructive stories for use in my school district's human sexuality unit, and I have used stories to help teach history and science. It seems appropriate to continue this tradition of using stories in my investigation of students' test taking experience.

Just because I am familiar and comfortable with stories did not mean my inquiry was natural or easy. Josselson, Lieblich and McAdams (2003) say of narrative inquiry,

In considering people as constructors of their experience, such research takes a giant step away from parsing human experience in predefined variables and requires of the researcher an equally major shift in perspective and approach. Rather than forming hypotheses, the researcher frames questions for exploration; in place of measurement are the challenges of listening deeply to others; and instead of statistics are the ambiguities of thoughtful analysis of texts (p. 276).

I believe this explanation of the researcher's role contained a personal challenge. I had to commit to listening deeply to my participants and I had to surrender myself to the uncertainties inherent in the process. Instead of pursuing the "why" so familiar to the quantitative researcher, I was seeking the "how" and the "what."

Research Setting

It is important for the reader to understand the demographic composition of the school district and school from which I drew my participants. The size of the school and district, the economic status of the student population, and the ethnic distribution distinguished my research setting and likely has influenced my findings. For this reason I have included demographic information here. In order to protect the anonymity of the participants, their school and their school district, I will refer to the Colorado community where the school is located as the fictitious city of Ute Creek and the school district as the Ute Creek Unified School District (UCUSC). The district contains few low-income and minority students and exceeds state and national means on both ACT and SAT college entrance examinations. UCUSC also exceeds state means on the Colorado State Assessment Program (CSAP) assessment. Additional descriptive data for Ute Creek Unified School District can be found in Table A ("School accountability reports").

Table A

Ute Creek Unified School District Demographic Information 2003-2004

Total students (K-12)	24,891
Students eligible for free or reduced lunch	27.8%
Special education students	9.8%
White students	79.3%

My interview participants were students from Silverheels High School located in the northwest portion of Ute Creek, Colorado. In 2004, the year prior to my participant interviews, the school received an overall academic ranking of “High” from the Colorado Department of Education based upon its performance on the Colorado State Assessment Program (CSAP) assessments. Additional descriptive data may be found in Table B ("School accountability reports").

Table B

Silverheels High School Demographic Information 2003-2004

Student enrollment (9-12)	1,871
Student dropouts	3.1%
Student enrollment stability	95.4%
Students eligible for free and reduced lunch	11.3%

Research Questions and Purposes

Rather than beginning with a list of research questions, I intended to allow data and experience to shape the questions I considered. I began with a central, research question: What did tenth grade students at Silverheels High School experience as they took the Colorado State Assessment? To begin answering this question with students I asked some very broad introductory questions such as, “Tell me about your elementary CSAP testing experience;” “Tell me about your testing experiences in Junior High;” and “Talk to me about how your peers view testing.” As I proceeded, other lines of

questioning were suggested by their responses. Anne Rogers (2002) recognizes the wisdom of taking this approach with qualitative studies. She states, “Researchable questions commonly become clear only after one has been involved in a research project for a considerable period” (p. 54).

Along with this central question, three research purposes guided my exploration. These reflected my intentions, but did not bind me to specific lines of questioning or methods of analysis. My purposes were:

- To solicit, listen to, and tell student stories about high-stakes testing.
- To apply appropriate and meaningful analysis strategies to student narratives.
- To explore links between student narratives and appropriate theory in education, psychology and leadership.

Selection of Participants

Josselson, Lieblich, and McAdams (2003) are averse to identifying a specific number of participants to be included in a narrative study. They state a study should continue until “saturation” is achieved. They admit this point may never be reached because it is likely the researcher will feel saturated before the data becomes redundant. Instead of an absolute number, they advise their students to specify a range for the number of participants. Heeding their counsel, I planned to include five to seven participants in my investigation. To achieve this number of participants I initially contacted eight students. Six of these students chose to participate, were interviewed, and continued in my study.

My sample was a convenience sample drawn from the sophomore class at Silverheels High School (SHS). The counseling staff at the school agreed to help me

generate a pool of participants. Because students needed to be verbal so I could generate a narrative, I asked the counselors to refer students to me whom they believed would be willing to talk. A second criterion I specified was that the participants must have taken the 2005 CSAP examination.

I conducted my interviews in the office of one of the counselors on the campus of SHS. The interior of the office was visible from the outside because the door contained a large window and there was also a window adjacent to the door. The door remained shut during the interviews to avoid noises from outside the office from distracting the participants. The interviews were held at times that were convenient for the students, usually during one of their free periods.

I first met with the students in early March, following the students' spring break. Prior to this initial contact I sent an informational letter to the students and their parents. This letter explained the objectives and procedures used in my study. During the first contact with prospective participants, I explained the nature of my study; detailed my goals, and defined the roles of participants and researcher. I also spent some time engaged in informal conversation with the participants to begin establishing a relationship. The purpose of this contact was to put the participants at ease and to introduce the study, not to begin the interview process. At the conclusion of these meetings I provided students with a formal consent form to be signed by them and their parents.

In the spirit of narrative inquiry the interviews were guided by my questions, but not bounded by them. During the interviews my goal was to learn about students' prior testing experience and about their impressions regarding the influence of testing on

school curriculum. In addition to the primary interview questions suggested by my research questions, I asked probing questions regarding their individual experiences. For example, I responded to a student's answer in the following manner, "You mentioned you didn't mind taking the CSAP exams as an elementary student, but you now resent taking them. What is different for you now?" In addition to probing questions, I asked some questions that allowed the respondents to step away from their personal experiences. Chase (2003) suggests using questions that invite the interviewees to theorize about people's general experience. For example, instead of asking the student, "Do you have difficulty with the CSAP exams?" I asked, "Do you think some students struggle with the CSAP exam?" and "What do you think is difficult for them?" Since I was collecting information for personal narratives, I needed to guide the discussion back to their own experience, but I hoped this exercise would help the participants understand their own experience.

I scheduled two interviews with each participant. The second interviews were conducted within two weeks of the first. In these interviews I followed up on some of the responses from their previous interview and explored their thoughts about their testing experience. It was my hope that this most recent, and their final, CSAP testing experience evoked strong reactions. I was particularly interested in knowing if the process of discussing testing had any effect upon how they viewed their testing experience. Did talking about the test and relating their testing experience bring about a change in the way they perceived the test?

Following my preliminary analysis, and after I had written narratives, I provided the participants with drafts of the narratives so they could review them. This exchange

was done via e-mail or with the help of a school counselor. Connelly and Clandinin (1990) emphasize the importance of this member-checking procedure. Not only does it contribute to the overall trustworthiness of the study, it also gives participants an opportunity to be an active participant in the research process. Connelly and Clandinin also stress the narrative process is to be empowering for the participants. This confirmation procedure provides a way to encourage such empowerment.

All of my interview sessions after the initial introductory contact were tape-recorded and transcribed. The interview transcripts served as data sources for my analysis.

Data Analysis

Josselson, Lieblich, and McAdams (2003) caution against choosing a method for analyzing narrative data prior to completing a study. They view qualitative analysis as a craft, and suggest that the use of off-the-shelf methodologies may not be successful. They are not alone in cautioning against prescriptive analytic methods. Connelly and Clandinin (1990) also avoid giving specific strategies for analyzing transcripts. They do stress the importance of hours of reading and rereading. Once the researcher is quite familiar with the data, they recommend using literary terms to guide the identification of themes that emerge as a result of this analysis. Scene, plot, character, tension, end point, narrator, context and tone are some of the structural components they suggest should guide an analysis.

Josselson, Lieblich, and McAdams (2003) are no more specific in their advice regarding analysis of narrative data, saying, "How this is done is the art of this kind of work and the process cannot be specified a priori" (p. 268). Like Connelly and Clandinin

(1990) they stress multiple reading and rereading during the process of identifying themes. When they work with their graduate students on a dissertation study they recommend students submit a demonstration interview and analysis. They believe results are more important than process. The pilot gives them an idea whether the student's methods will yield a satisfactory result.

As Josselson and Lieblich (2003) and Connelly and Clandinin (1990) suggested I found an analysis strategy as I worked through the data. I followed the narrative writing procedure I had anticipated in my proposal but the remainder of my analysis followed quite a different path.

Before beginning to write the individual narratives, I read and reread the transcriptions of the interviews. On the first pass I read the transcripts without making any notations. This initial reading was just to familiarize me with the data. On the second and third passes I made notes in the margins and listed thematic elements on a separate sheet. I focused on one participant at a time, beginning the writing as I referred to the notes and transcript text. I refined the narratives and then left them as I moved to begin the next narrative. I returned to the first narrative for final revisions after I had finished the other five. After I had revised them to my satisfaction I sent them to the participants for member checking. Based upon the suggestion of one participant I changed a single word in his narrative. The other two participants made no suggestions for change.

The remainder of my analysis was not as straight forward. As I explored different analysis strategies I was drawn to a process explained by Clarke (2005). She developed a method she calls Situational Analysis for use in grounded theory studies. I thought it would work well with my narrative investigation. As suggested by Clarke, I began by

constructing a messy/working version of an abstract situational map. This map included pertinent human, non-human, material and symbolic/discursive elements present in the analyzed data. I later refined the categories, using some suggested by Clarke and others suggested by the data to create what Clarke describes as an ordered/working version of the situational map.

Utilizing the situational categories as codes, I started a line-by-line reading and sorting activity using NVivo software. As I worked towards completing the second participant's transcripts I began to question the value of using this analysis strategy. I was worried that although the categories that I was using for coding provided a means of sorting data, they ultimately would not provide me with any insights regarding important themes and commonalities among my participants. Despite my reservations, I persisted.

I had nearly finished analysis of the second transcript when I suffered a loss of data as I was upgrading a computer program. I had neglected to sufficiently back-up my data so nearly all of my coding was lost. Frustrated, I ceased working on the analysis for over two months.

I still had the intention of pursuing the same analysis strategy when I revisited the transcripts and narratives after my hiatus. As I reread the data, themes jumped out at me with surprising clarity. I had been thinking about some of these thematic elements as I had written the narratives and new ones emerged as well. It occurred to me I now had a better way of analyzing the data than returning to the situational mapping strategy I had used before.

With the support of my methodologist, Jim Banning, I embarked on a new analysis strategy. Using themes I had identified, I went back to the transcripts again,

lifting quotations that supported each thematic element. This was a way for me to confirm that the themes were indeed present in the text and not just in my mind. I merged some themes and renamed others but most remained similar to when I had first identified them. The thematic elements that I identified in this process are presented in Chapter Four

I have included my false start with situational analysis here because I think it illustrates the difficulty a researcher using narrative inquiry may have identifying an analysis strategy prior to collecting and reading the data. Creswell (1998) emphasizes that familiarizing oneself with the data through reading and rereading is critical to the analysis. This is what I was doing even as I was pursuing the situational analysis strategy I later abandoned. Even as I proceeded down this unproductive path I was becoming familiar with the data by exploring some thematic elements I later used in my final analysis.

Ethics

My fifteen and sixteen year-old participants were vulnerable in many of the same ways younger children are. It was my responsibility to guide their story telling without guiding their thinking. If their opinions changed as a result of telling their stories so be it but if their opinions changed because I imposed my beliefs upon them that would constitute an ethical dilemma. This meant that as I conducted my interviews I needed to try to encourage their telling without layering my beliefs over theirs. I attempted to be an active, but not imposing, questioner.

To gauge whether I had influenced the attitudes of my participants I asked them at the end of my second interview what they thought my opinion regarding CSAP exams

was. None thought I held a critical view of the process and two of the participants thought I was supportive of the testing program. From these responses I conclude that I did not telegraph my opinions regarding high-stakes testing to the participants in my study.

Other ethical issues arose during my data analysis and writing phases. I had an obligation to thoroughly analyze the data, and once the data were analyzed, to take care to avoid distortions during the narrative writing stage. Clandinin and Connelly (2000) warn of a tendency of narrative writers to smooth the writing. That is to say the writer sugarcoats the content to make it appear less harsh or more positive. The authors suggest a narrative researcher should reveal any selections he or she has made that may alter the text. They also believe it is important that the researcher adopts a critical role at the same time he or she functions as the storyteller. This is good advice and it was important to consider early in the development process. I would like to think the influence of my empathy or antipathies towards participants did not affect my narrative writing but I cannot rule out the possibility that they may have had an effect.

Trustworthiness

Ultimately, both credibility and transferability relate to the care and thoroughness with which I conducted my study. To ensure this was true I attempted to conduct my interviews in a way that encouraged story elements to emerge. It was also necessary to exercise care in analysis by engaging in the “relentless reading and rereading” discussed by Clandinin and Connelly (2000). Creswell (1998) recommends using at least two strategies from a list of eight he says increase the trustworthiness of a study. One strategy that I used and have previously mentioned was member checking. In my study this took

the form of asking my participants to read the narrative drafts of their own stories so that they could comment regarding their accuracy. Another Creswell strategy is clarifying researcher bias. I have already added a statement revealing my bias towards high-stakes testing in the introductory chapter.

Personal Statement

I am a middle-aged, middle-class white male with 25 years of teaching experience. Twenty of those years were spent in one elementary school. When I began teaching in this building in the early 1980s most of my students walked to school from the upper middle-class, predominately White neighborhood on the south side of the school. Fewer students came from the low to moderately priced rental duplexes and apartments to the west and north of the school. As the neighborhood matured, fewer school-age students came from the affluent neighborhood to the south and more students came from the less affluent and more diverse neighborhoods to the north and west. This meant that as my teaching experience increased, so did my exposure to low-income and minority students. I am currently teaching in a school with a similar demographic composition. Many elementary schools in Ute Creek see little economic or racial diversity. The demographics in the two buildings where I have spent most of my career, while not as diverse as those of a typical urban school, represent more diversity than is found in many schools in this community. Prior to teaching in the Poudre School District I worked as a teacher's aide in a Denver school with a large number of students of color and from low-income families. I later substitute taught in numerous inner city schools in the Denver Public Schools. These experiences with middle class, minority and low-

income should have been sufficient to prepare me for the limited diversity I encountered among the participants in this study.

My teaching experience contributed to my preparation for this study in other ways as well. I have prepared students for accountability testing, I have given high-stakes tests, I have watched students react to tests and test results, and I have participated in planning sessions where test preparation strategies have been discussed. I have an insider's view of this process, but only for the testing of elementary age students. My understanding of the process may have assisted me in asking more directed questions about the students' testing experience. My lack of experience with high school students could have had the opposite effect but I was pleased to find that I was quite comfortable interviewing the older students. They seemed at ease with me as well.

I am a person of privilege who interviewed at least one student with a background of poverty. I am a teacher who interviewed some students who may have had bad experiences with other educators. I am a person who is familiar with the testing system who needed to see it anew through the eyes of students. I am an elementary school teacher interviewing high school students. Any of these personal characteristics may have interfered with my effectiveness as an investigator. It is my hope that by acknowledging these contradictions that I was be better able to recognize the influence my experiences and biases may have upon my study.

CHAPTER 4: RESULTS

Introduction

I begin this chapter with a brief description of my participant population. I then provide a detailed narrative describing each of the participants. Next, I present seven thematic elements identified as a result of data analysis. I then explore these themes in relationship to the individual participants. The chapter concludes with brief summative statements.

Participant Overview

The six participants in my study were all high school sophomores. I interviewed them soon after they completed their last CSAP exams in April and May of 2006. Michael, Thomas and Dennis were in the International Baccalaureate (IB) program and Jackie had recently transferred out of IB to take Advanced Placement (AP) classes. Audrey and Maria were taking less rigorous standard level courses and were receiving resource support, likely because they had been identified as having learning disabilities.. Four of the participants were White and two were Hispanic. All volunteered to participate in the study after meeting with me to go over the research protocol. Following the introductory meeting, the six students met with me for two taped interviews that lasted approximately one-hour each.

Narratives

I developed the following narratives over a five-month period that involved extensive reading of the interview transcripts, writing, revising and reflection. My purpose was to capture the participants' stories so as to gain a better understanding of how students experience, understand and interact with the CSAP exams. As I read and reread the interview transcripts I made notes in the margins and recorded dominant thematic elements. I also referred to my researcher's notes, recorded during and soon after the interviews. I wrote my initial drafts of the narratives using the transcripts and notes. This allowed me to acknowledge non-verbal clues provided by the participants such as body language, tone of voice and facial expressions. I offer these narratives as portraits of the participants. Like a visual portrait, they are a synthesis of the creator's perspective and the images presented during the shared encounter. Following each of the narratives are some short vignettes. Their appearance in italicized font signifies a change in perspective.

Most of the content in the narratives is devoted to a general description of the participant's experiences and thoughts. These portions cover broad swatches of the participant's life. The image presented by this description could be compared to those created using a wide-angle view of a camera lens. The vignettes signal a narrowing of focus. Single events in the participant's life, or in one case, a participant's testing fantasy, are explored in these passages. This technique applies an analysis strategy suggested by Pamphilon (1999) to guide interpretation of personal narratives. When a zoom lens is used with a camera it allows the photographer to vary the focus of the lens. By adjusting the focal length of the lens the photographer can view either a single flower or a

panoramic landscape. When this metaphor is applied to narrative analysis it enables the viewer to shift focus and acknowledge both foreground and background elements within a narrative.

Pamphilon developed this strategy while examining positive aging among nine Australian women. She realized that social, historical and personal influences affected participants' life events. By zooming in and zooming out she was able to shift her focus to include these various elements and this enabled her to better understand, and consequently explain, their narratives. I have used the technique in a similar way to move from a more holistic view to a more focused view of my participants' stories. In the vignettes I zoomed in on the thoughts and feelings my participants revealed to me regarding a single event in their lives.

Narrative #1: Thomas

Initial Impressions/ Description

As I rose to meet Thomas at our first meeting I was struck by his imposing size and confidence. He was a big fifteen year-old with powerful shoulders and arms. He shook my hand firmly and with conviction, looking me directly in the eye. While other participants required some time to relax and open up, Thomas was at ease and ready to start the questioning process as soon as he sat down at our first meeting.

Background/ Interests

Thomas came to both interviews directly from swim practice. Swimming is just one of many extracurricular activities in which Thomas participated. He played soccer and baseball in elementary school, played on a hockey team for several years and

frequently played tennis with his mother. In addition to his athletic pursuits he sang in the school choir.

Thomas is an athlete, but he did not identify himself solely in these terms. From an early age he found pleasure spending time with students who had sports as their focus as well as with those who valued academics. He found himself moving between two peer groups, one consisting of students with athletics as a focus and the other directed towards academic pursuits. “I think kids in both groups kind of recognized that we could be in both groups. It would be nice for the kids to be able to see that there’s not much of a difference and that people can be involved and it’s fine.”

Testing History

Thomas knew that some of his peers disliked and struggled with the CSAP exam but it was hard for him to understand what they experienced. He took the tests confident in his abilities and sure that the outcome would be good. Generally his confidence was justified; his performance matched his expectations. The CSAP assigns levels of proficiency based upon the students’ performance. These levels, from lowest to highest are, unsatisfactory, partially proficient, proficient or advanced. With few exceptions Thomas achieved performance designations of “advanced” on the exams. The few times he did not, his scores fell within the “proficient” category. Because of his testing success it was difficult for him to understand why it was a struggle for others. “It’s not a huge deal for somebody to sit down and take the CSAP tests. It’s not too difficult because you just rely on what your teacher will tell you and do your best.”

Testing Reflections

Being extremely goal-oriented, Thomas wanted to know what content was going to be on the CSAP exams. He was not opposed to working hard, to studying for the test, but he wanted to know just what it was that he was to be accountable for. When I asked him what message he would want to send to legislators regarding the test, he responded by asking for more direction. “So maybe if they could possibly, you know, just give general topics to teachers so that teachers could plan their lessons around those general areas as well as their regular curriculum.” Just like he knew the goal in hockey was to gain the most points and win, his goal on the CSAP was to know as much about the topic as he could so he could get a good score.

When Thomas discussed peers who had chosen not to take the CSAP exam, his strong work ethic and his loyalty towards school showed. “I feel like everyone should take the test, and you know, even if you don’t feel like that you’re going to do extremely well, you should just duke it out and do the best you can.” And you know, even if they can manage a 50% on it it’s better than absolute zero on it.”

Despite being a model student, Thomas echoed the opinion of some of my less eager participants regarding test accountability and relevance. When discussing the kind of results he expected and received on the test, he first mentioned his successes but then he summarized stating, “It didn’t really matter to me how well I did or not. Because it wasn’t like a grade that went on the report card or anything.”

Peer Responses

Hard as it is for him to relate, Thomas recognized that some of his peers struggled with the assessments. He mentioned two students who did not take the CSAP exam

during his last testing sessions. A friend decided not to take the science exam because she feared she didn't know the material and would fail anyway. Another student, a classmate from his English and chemistry classes, was able to convince his parents to exempt him from the CSAP testing.

Insights/Unique Perspectives

Thomas thought his athletic background had helped him with academics and with test taking because it had taught him to keep his goal in sight. He talked about the importance of having focus. "In sports you always have a goal, you know, to win as many games as you can, to win the championship. Just to have a goal and to be able to work on that in every single practice, that carries over into academic life." Thomas also recognized how participating in team sports may have contributed to his academic success. "You know, hockey is my main sport right now and you have to be able to combine and work with all of your peers toward that one goal. You just push yourself to be all that you can."

Conclusion

The Thomas I interviewed seemed to breathe confidence with every breath. He did well in school and he has done well in the athletic endeavors he has pursued. His parents have inculcated him with an understanding of the importance of working hard in school and they have impressed upon him the necessity of having a default position. Maybe he will be an Olympic swimmer or maybe he will be good enough to be a professional hockey player, they have told him, but he better not count on it. He will have to work hard at his schoolwork too if he plans to have a successful future. Successes have followed previous successes to the point that Thomas has come to expect success.

Success breeds confidence and confidence has led him to enter activities with sureness, a clarity that many of his peers lack. The CSAP exam gave Thomas just one more opportunity to fix his gaze upon a goal, to push his pencil with the same conviction with which he swings a tennis racket. If he prepared well enough, if he focused hard enough, his test scores would match his expectations.

Vignettes

The hockey stick.

Thomas: Dad, check this out, my hockey stick is cracked.

Dad: (Inspecting the stick) I think if we tape it you will make it through tomorrow's game, but it looks like you will need a new stick.

Thomas: You remember that one we saw last week at the Ice House?

Dad: I sure remember the price. If you want that one you will have to buy it yourself, or at least pay the difference between it and another one like this one."

Thomas: That's what I thought you'd say, but last week you and Mom bought me that expensive graphing calculator.

Dad: Yes, we bought it for you, but that was for school, and that is altogether a different story. If it is for your education we will make sure you get whatever you need, even if it is costly. For your sports activities we are happy to provide you with the basic stuff you need. If you want something better you'll have to pay for it yourself. It looks to me like you have been doing pretty well with this stick; didn't I see you score a goal with it just a few minutes ago?

Thomas: I'll think about whether I want to spend the extra money for a better one.

Total immersion.

Thomas enters the testing room with the same confidence he demonstrates as he steps onto the starting block at a swim meet. He is rested, mentally prepared, and certain that he will do well. Just as the cool water engulfs him as he dives into the swimming pool, shutting out most sensory input, the test taking process closes around him as he begins working on his CSAP exam. He scarcely notices the other students taking the exam in the room with him. They are as distant as a swimmer three lanes away when he is competing in a 100-meter freestyle race. And just as he would in the race, he focuses absolutely on the task at hand and pulls hard to the finish.

Homework.

Thomas arrives home from school, grabs a snack, calls his mom to check in, and sits down to do his homework. Even though his parents aren't there to remind him about getting his work done, he works diligently to finish his math and English assignments. As he returns to the table from the refrigerator with a glass of milk in hand he glances through the window and sees some neighbor boys rollerblade past. A tinge of regret grips him momentarily but it quickly passes as he sits down and takes up his calculator.

By the time his work is done his parents are home and dinner is cooking. Now that darkness has fallen his buddies are inside as well. His homework complete, he watches television for a few minutes before his family sits down to eat.

After dinner his dad suggests that they take a look at his homework assignments to make sure that the math problems have been completed correctly and that his English assignment doesn't have any grammatical or spelling errors.

Narrative #2: Dennis

Initial Impressions/ Description

Tall and thin with tightly curled blonde hair, Dennis was bright, thoughtful and eager to talk to me. It was apparent that Dennis was a thinker. He contemplated each question I posed and replied with carefully measured responses. He had a keen desire to appear neither arrogant nor inaccurate in his assessments of his school experiences. A few times during the interviews he revised statements after considering their authenticity or when he felt he was being too boastful.

Twice, when discussing his CSAP performance during his first interview, he referred to himself as an "outlier." This puzzled me, so in the follow-up interview I asked him what he meant by this. He first told me he used the term to describe himself because he had scores on some of the CSAP exams in the upper percentile range. This would mean that he was an "outlier" because he was an elite performer. After some thought, he amended that characterization to describe his nonchalance about taking the test. "I perceive the test a bit differently and I don't take it so seriously," he explained. In this sense he thought he was an "outlier" because he did not perceive the test to be as important as his peers.

Background/ Interests

The label, “outlier” may be applicable to more than just Dennis’s test performance, I think it also describes how he has responded to his school. Dennis first theorized that his disconnect with school might have something to do with the number of schools he had attended. He went to three different elementary schools but only a single junior high and high school. Although this may be more schools than many of his colleagues have attended, it is not an extreme example of transience. As a teacher, I have had many students who have attended more schools than Dennis. Furthermore, as he pointed out after considering it for a while, some of these changes were natural transitions from one schooling level to another. Dennis’s relationship to school may have more to do with the way he perceives himself and relates to peers.

Dennis mentioned twice during our interviews that he had few friends during his early educational years. He said he found some companionship in his gifted and talented classes but felt little closeness to the rest of his peers. This lack of connection may have led to a lack of loyalty to his peers or schools. Despite his disinterest in schools, he mentioned he felt a bond with some of his teachers. These bonds, however, did not necessarily translate into a positive connection with the school itself.

When discussing his feelings about Silverheels High, Dennis said he felt he belonged among his fellow IB students. He thought he was in the best educational setting, yet he did not have a strong connection to the school. He attended classes, did well academically and connected with some of the other IB students, but he seemed to be holding back, or, perhaps he was just reluctant to make the sort of emotional plunge that was required to become truly connected with school. Perhaps it was because he had been

in several different schools and had protected himself from hurt and disappointment by holding school at a distance. Maybe it is more than that, perhaps this outlier liked maintaining a little distance and preferred watching from the margins, somewhat like a guest at a dance party who watches others dance but does not participate.

Testing history

Whether it was because of his frequent moves, or because he had difficulty establishing friendships, Dennis did not form strong connections with his elementary and junior high schools. This attitude carried over to how he felt about Silverheels. Dennis mentioned in our conversations that he did not feel a strong loyalty towards Silverheels. Although he thought the IB program was an excellent fit for him, and even though he said he appreciated his teachers at SHS, he did not feel compelled to perform well on his tests for the sake of the school. It appears from our conversations that he generally did well on the tests but that his performance was driven by a personal desire to do well rather than by any sense that he must perform to benefit his school.

“When I was in elementary school people always said I was a self-motivated. It’s a matter of having to care enough about the test to do your best,” observed Dennis. This motivation did not extend to all areas of testing, however. His handwriting and spelling were poor according to his own appraisal, so he was not particularly motivated to complete tasks that required him to write or display his poor spelling abilities.

Motivated or not, Dennis confessed that taking the CSAPs was a stressful experience for him. He told me he had fallen asleep following a testing session. “I never get enough sleep. I don’t usually sleep in other classes, it’s because I’m presented with the opportunity and I did something that I really didn’t want to do, and so it kind of tired

me out,” he remarked. Other students told me they had either slept following finishing their exam or had observed other students sleeping afterwards. Dennis is the only participant who attempted to analyze this phenomenon. Other students attributed this sleeping solely to boredom. Dennis’s explanation seems more credible.

Testing Reflections

Dennis mentioned that he was a willing and capable taker of CSAP tests when he was in elementary school. As he progressed through school and his thinking became more sophisticated he became a testing critic. His biggest objection to the test was that he believed that it was designed to target students who performed in the center of the bell distribution curve. Dennis felt since he was two years ahead in math and had forgotten some of the formulas and algorithms needed for the test, that it did not accurately measure his true mathematical ability. Another problem Dennis had with the test was with the writing prompts. Dennis wished he could have written authentic responses to the exam questions. He wanted to see more interesting prompts and he would have liked more time to work on the writing exercises, including more time to revise his finished product. He thought there should have been more opportunities to include research in the testing process. To further mimic authentic writing tasks, Dennis thought students taking the CSAP exam should be able to use computers to do their writing.

Dennis had little faith in the system used to evaluate student responses on the portions of the CSAP exam. “When you have something that they bubble in the answers and you have one of four answers, it’s not a really good measurement of a person’s understanding of something like poetry.” Dennis wished the test would get at the thinking behind a written response. He felt the test looked more at whether certain conventions

were in place or whether the student could identify a literary device such as a metaphor or a simile than at a deep understanding of the topic.

As the reader can discern Dennis was critical of many aspects of the CSAP testing process. Some of his other issues with the exam were related to control. The CSAP testing environment was precisely defined and carefully controlled, much more so than when he was given a test by one of his classroom teachers. Dennis described the difference in the following response:

Yeah, my English teacher gives me a test, first off I know the person that's going to be reading it and I'm writing to them rather than at this piece of paper. I love it. First off you could make a noise and they wouldn't invalidate the test.... It's a very weird thing this testing experience because it's not like anything else you experience at school; it's very serious.

For Dennis, a connection to the teacher and a little bit of flexibility goes a long way towards making a testing experience tolerable.

Another control issue that irritated Dennis was the enforced silence during CSAP testing. I asked five participants about their relationship to silence during testing and about silence in their broader lives. Like the others, Dennis said he was not comfortable with silence during the exam, but he wisely linked his dislike to control. "In any other situation you would be able to choose it... You don't matter, it's not you, it's the whole," he stated. Just as Dennis resented the lack of the personal in his responses, he interpreted the enforced silence as another means of managing the testing process, of making it easier for anonymous authorities.

Peer Responses

While he counted himself among students who were motivated to do well on the CSAP exams, he knew that many of his peers were less enthusiastic. He was not sure that

external forms of motivation were effective in overcoming this lack of enthusiasm. He was skeptical of incentives schools used to try to motivate students to do well. “I don’t think if they don’t care in the first place that you’re going to be able to motivate them at all,” he remarked. When questioned about what specifically motivates students, Dennis mentioned family attitudes regarding education as motivators but he also talked about how motivation follows success. The implication is that his less successful peers would be less motivated to do well on the test.

Insights/Unique Perspectives

Dennis had a novel suggestion for test designers. He thought that all students taking the test should be provided with a common experience, which could then be used as a departure point for writing their essays. In this way the playing field would be leveled somewhat. He reasoned that students with limited life experiences would not be at a disadvantage if this system were used.

An assessment should be a learning experience Dennis believed. It bothered him that he had gotten no specific feedback from his CSAP exams. It takes several months following the exam before students receive their scores. When his scores did arrive the scores gave little feedback that Dennis valued. They put him in a performance category: advanced, proficient, partially proficient, but they gave little direct information about how he could improve his scores on future tests.

Whenever we discussed the CSAP, Dennis bristled. It was clear that he did not like being boxed in by restraints posed by the test or the testing process. When I asked him what the ideal testing situation would look like, he suggested sitting one-on-one with an evaluator so that they could find out who he really was. He challenged the test

designers to look at him as an individual, to recognize his uniqueness and to allow him to express that uniqueness.

Conclusion

Dennis is a social critic. He thought deeply about his relationship to school and about the relationship between testing and learning. His keen perceptions regarding his school environment set him apart and put him in the role of an interested but dispassionate observer. His “outlier” status positioned him well to appraise the test from a perspective more removed than that of his peers.

Vignettes

A testing fantasy.

“Hi Dennis. I am your personal CSAP evaluator. I will be spending the next two weeks with you so the state of Colorado can get a complete and comprehensive view of your abilities in multiple discipline areas. The first step is to make you comfortable with the process. I will ask you some questions about yourself and you may ask me questions about the evaluation process. After I find more about you, I will be giving you some meaningful tasks to complete that are tailored to your ability and interests. You will have access to the school library, a computer and the Internet as you complete these assessment tasks. You may notice that I will be occasionally taking notes. I will be making observations as you go about your work. This information will give us additional information about your abilities. Don't be stressed by this process. Remember, we are committed to learning about who you are as an individual. We will work with you so that you can demonstrate your unique strengths. Are you ready to begin?”

Another reminder of the same inadequacy.

Dennis is taking the writing portion of the CSAP exam in the library with about 40 other students. The principal, who is proctoring the exam, briefly stops behind his chair, and then moves on. Dennis becomes absorbed in his writing task and loses track of where the principal is in the library. He has nearly forgotten about him when he notices his return. Before, he had stood fairly unobtrusively behind him. Now he is at his side, openly examining his writing and, to Dennis's surprise, he speaks. In a whispered tone he asks, "Can't you try to make that a bit more legible?" There it is again, since his early elementary grades he has heard teachers complain about his handwriting. "You have good ideas, I just can't always make them out," was one of the kinder remarks he had heard from them. Others, such as this most recent rebuff, are more pointed and more difficult to accept. He often tried to make the perfectly shaped letters that some of his classmates produced, but somehow the lines were never true, the sizes of the letters never consistent. "I'll try to make it more clear," he mutters without looking up. To his relief the principal moves on.

Narrative #3: Michael

Initial impressions/ Description

Michael was an earnest young man who worked hard at school and enjoyed the support of his family. He was an eager participant in my study. "If anybody is going to ask my opinion, I'm going to give it to you," he declared during his first interview. Besides being eager and ready to state his opinion, Michael also professed a strong work ethic. He told me "If you're going to do something, do your best. You're just wasting

your time if you don't." This credo, whether original to Michael or a product of his upbringing, was a guiding compass for him. He restated this belief several times during our interviews. Later in the interview process he referenced his parents' opinions when discussing the CSAP exam, saying they thought he should take the exam and do his best because the school would be penalized if he did not take it. Because this belief strongly resembles his own declaration, I suspect the seed for its formation originated with his family's value system.

Background/ Interests

During Michael's first few years of schooling he was living with his family in New York where he lived in a small community in the far north of the state. According to Michael, the smallness of his school ensured everyone knew each other. He mentioned because of the size and intimacy of the school there were no cliques. Everyone knew each other and supported each other. "It was more like family," he said.

This was the same feeling that Michael experienced among his peers at Silverheels. It was his opinion that like his small school in New York, Silverheel's climate prevented the formation of cliques. It was not that there were not groups of students who congregated with each other he explained, it was just that these groups did not interfere with his freedom to move freely between them. Michael found many ways to connect with his peers at school. He described himself as "very active," and his list of activities supported his claim. He played on the tennis team, managed the girl's tennis team, played basketball during his lunch break, played in the band and hung around with his fellow IB students. He planned to participate in the National Honor Society as a junior and he was recruited to participate on the Science Olympiad team.

Although Michael was quite involved in his school, his involvement was a new thing for him. Whereas most of Michael's friends enjoyed the company of peers throughout their school years, Michael had just rediscovered the spirit of community he had experienced in his small elementary school in New York. He was again connected with students who shared his interests and abilities.

Testing History

Michael was not an eager writer. He confessed that although he could write an IB essay upon demand, writing personal responses as required by the CSAP was not a pleasurable experience for him. "Whenever there is writing on anything, I kind of moan inside. And I guess just for an instant it might slow me down but I just do it and get it over with. It's got to be done," he told me. This statement reveals a mixture of anxiety and determination.

To make the writing task easier he relied upon a writing format he learned in English classes called a power paragraph. The structure is predictable and easily followed. First the student writes a topic sentence, often containing a number reference i.e. "I like to write for three reasons." In the subsequent sentences the three reasons are enumerated. The final sentence in the paragraph summarizes or restates the topic sentence. This form taught to elementary students as early as the second grade provides a secure, if not creative way of approaching an academic writing assignment. Michael recognized that this was not the most eloquent way to approach CSAP essays but it was certainly an expedient and safe approach.

At first he said that if he could write about a less expository topic, such as the current immigration debate, he would be able to do so more readily. When I probed a

little more he confessed that he was generally not comfortable with the writing process and that it mattered little what the topic of the essay might be. Michael had accepted that it was necessary for him to improve his writing. His dad was an English teacher and had offered to assist Michael with this task. He indicated to me that he was open to receiving some tutoring from his dad over the summer.

Beyond the writing issue, the test-taking process itself was unpleasant to Michael, “Boring” he called it, but worse was the anticipation of the test. “... like it’s really not anticipation but like the dread of having it come up; you don’t really want to take it,” he explained. This opinion was reinforced when he told me how elated he was to have finished his last CSAP exam, knowing that he would never have to take one again. Clearly, he was glad to have this experience behind him.

Testing Reflections

As a compliant test taker Michael appreciated the test preparation that he had received from his teachers. He did not express any discontent about the time he had spent preparing for the exams. He said he thought those preparations had helped reduce his anxiety about taking tests. One of his few complaints was that his math teachers did not prepare him sufficiently for the math CSAP. Like other participants, he expressed discontent that his current math instructional level was above that targeted by the test. These students said this put them at a disadvantage when taking the test because it had been two years since they had been accountable for the math content on the exams.

Although Michael was serious about his test taking, he had no illusions about what the results said about him as a person. “What do you think the test says about who you are?” I asked him. “Personally? Who I am? Nothing,” he replied. He went on to say

that his test scores only told about how well he takes tests. He explained that the tests just displayed how well he remembered certain content. He was certainly not one to harness his opinion of himself to the test scores. The tests, in his opinion, only told about one small part of the person he was.

Michael was also uncertain about the relevance of the test. Like some other participants, he viewed the CSAP in a different light than the college entrance exams that he would face in the next two years of his schooling. He understood the purpose of those assessments. The relevance of the CSAP to his individual situation was less clear to him.

Peer Responses

Michael realized that his acceptance of the CSAP was not the norm. When I asked him what his classmates thought about the tests his first response was, “They think it is stupid.” He elaborated when I asked for more detail, “It was just another test. I mean we already are being tested so many different ways.” He explained it would be better to use data from tests they had already taken rather than forcing them to take another test. Although I asked for his peers’ opinions, I believe I received more information about how he felt about the CSAP exams.

Insights/Unique Perspectives

Michael wondered about the imposed silence during CSAP testing. He discussed some ideas he had for allowing students to enjoy music during their test taking experience. He suggested that students could be sorted according to their music preferences prior to the CSAP exam. Those who preferred rock and roll would be in one room and there would be other rooms for those who preferred hip-hop, country, or heavy metal. Although he recognized the impracticality of the solution, it demonstrated the

importance of music in his life. In this regard he was like the other students I interviewed. They have music available through their MP-3 players whenever it is permitted in their environment. It seems logical to them that they should have access to music when they are testing as well.

Conclusion

Michael was a highly principled youth with strong beliefs about working hard at everything he attempted to do. Still, not everything he attempted was easy for him. Reading and writing tasks were difficult and caused him some anxiety when he took the CSAP exam. It is admirable that he held on to his strong work ethic in the face of his difficulties. Michael was also a budding social critic with valuable insights about his school environment and particularly about high-stakes testing.

Vignettes

In the cocoon part 1.

“OK students, now that you have the test booklets and your answer sheets, I want you to spread out in the classroom so you are not near any other students.” Michael watches as his fourth grade classmates stand up and begin to drag their desks to various places around the room. He briefly considers staying where he is since nearly everyone around him is moving but then he changes his mind. He stands up, drags his desk over to the coat racks and positions it so the coats drape around him. It is a little bit like being in a tent. In here he cannot see his fellow students and they cannot see him. He is safely cocooned within a fiber retreat.

In the cocoon part 2.

The room is quiet, unnaturally quiet. Most of the tenth graders are busy writing responses to the supplied prompt. Some students are frowning, one girl chews on her pencil, and a dark-haired boy looks at the teacher in the front of the room as if seeking inspiration. Michael is staring, his gaze unfocused. He stares for several minutes giving no clue as to what he might be thinking, seemingly oblivious to the passage of time. Suddenly he starts, moves his head, blinks his eyes, and rolls his shoulders slightly. He checks the writing prompt and the paragraph he has just written. He twirls his pencil several times. Now it is obvious that he is back thinking about the task at hand. After a few more minutes his pencil stops its slow spiral and he once again employs it as a writing instrument. At no point during the previous 15 minutes has he looked at his classmates, he has successfully sequestered himself within this classroom full of students.

Narrative #4: Jackie

Initial Impressions/Description

When I met Jackie I was immediately impressed by her confidence. She was more formally dressed for both interviews than other participants and her hair was carefully coiffed. She was a year younger than her peers and small in stature but she had the confidence and presence that many adults never achieve. I soon learned that she was also a high achieving student. From a young age her parents had communicated high academic expectations to her and her two brothers. She watched her two older brothers excel in school and listened to her parents express their desire that they do well. She took that message to heart and in her early elementary years, when Jackie was attending

school in a neighboring community, her teachers and parents made the decision to advance her a grade.

Background/ Interests

When she was in fourth grade her father's job took their family to London, England. She experienced some culture shock there but enjoyed being in small classes. She returned the next year to Ute Creek where her neighborhood elementary school had recently adopted the International Baccalaureate (IB) curriculum. Following elementary school and against her desires, her parents sent her to one of two junior high schools in Ute Creek that also had an IB program. As a freshman she began the IB program at Silverheels High School. During her sophomore year she made the decision to move out of IB.

In part, Jackie's decision to leave IB was driven by her desire to become active in student government. The IB schedule was so intensive that it left little time for her to serve on student council. Other factors in her decision were her desire to take some electives outside of IB and the lure of AP classes. More colleges, she said, recognize AP classes than IB courses.

Jackie had always been a popular child with many friends. This popularity was welcome but also brought with it some challenges. She frequently struggled with what it meant to be popular. There were times during her schooling when she observed people from the popular crowd she was associated with behaving like bullies. Other times she thought the group was elitist when it excluded other students. Disturbed by these behaviors she discontinued her involvement and turned her attention to other students. During her sixth grade year she withdrew completely from group affiliation and spent

most of her time with a single close friend. Her ambiguity regarding membership in the popular group was reinforced when she noticed during junior high that the “popular group” did not hold academic performance in the same high regard that she did.

Reflecting upon her schooling and group affiliation, Jackie felt her favorite social circumstances were when she was part of the small learning community in her English primary classroom or as a student at Silverheels. She described the lack of cliques as a positive aspect in both of these environments. She liked to be popular but not if that meant being part of a group that is arrogant and exclusionary.

An active student, Jackie was involved in school government and athletics. Beginning when she was two Jackie had also been involved with dance. More recently she had been a member of a cheerleader group. She played soccer, was on the track team and at the urging of her active family, was considering participating in distance running.

Testing History

Jackie had always done well on tests. She was not one to take them lightly because she knew that most tests were tied to a grade, or perhaps more significantly, to how the teacher or her parents thought about her. Now that she was older and thinking more critically, Jackie was questioning whether the CSAP exams had anything to offer her.

Testing Reflections

Jackie did not choose to talk to me because she was a CSAP critic with a list of complaints about the exam but as I interviewed her, her critical attitudes regarding the test became clear. Early in our conversations she expressed anger that so much class time was spent on preparation for the exams. Although she was happy to spend time preparing

for tests that were closely tied to her future, she resented taking assessments that were not linked to her success. “So teaching to the CSAP, as we call it, is kind of annoying, a waste of class time, because it doesn’t really have any benefit, whereas if they were teaching for the ACT or SAT, we would be more motivated to do well,” she complained. It was clear to Jackie that the CSAP had little impact upon them personally. She knew the results mattered to the school but she could not see a link either between her current learning or her future success and the CSAP exams.

Jackie had a pretty good idea about where she was headed in life. Her choices in school reflected this focus. All her high school courses were tailored toward college admission. AP examinations and college entrance tests were on her junior year agenda. To ensure her best performance on these exams, she intended to use a study guide or take an online course. When she considered the significance of the CSAP in her life she examined it with college in mind. She asked the question “How does this help me?” She thought the results came too late and were not specific enough to help her improve her content knowledge and she believed the test had no direct bearing on her future. She knew the scores went on her transcript but she had been told colleges either did not pay attention to these scores or they assigned very little weight to them. The bottom line for Jackie was that there was little to gain from performing well on these exams.

Early in the interview process Jackie expressed dissatisfaction with her teachers for spending so much time, “teaching to the test.” Later she admitted she didn’t mind it when teachers were teaching to the test if that test was one of her course exams. “They really want you to do well; they always teach to the test. They always want their students to know that, you know, what will be on the test, so they’ll be prepared.” When teachers

taught to the test it reduced her anxiety because she knew what content would be on an exam. What really bothered her was teaching to a test she did not think was important. After talking about the topic with me for a while she had a realization – teachers taught content they knew was on the CSAPs because they wanted their students to perform well on the test. They wanted the students to do well because it reflected upon the school as well as their teaching performance. For a moment she understood her teachers' motivation but I did not get the impression that she let go of her frustration over wasted learning time used for test preparation.

Peer Responses

Jackie had a strong sense of loyalty to her school and that is likely the reason that she did not take the CSAP exams with an indifferent attitude. She knew poor test scores from students who did not take the test seriously would hurt the standing of her school. Because of her sense of commitment was so strong, it was hard for her to understand how someone could throw away a test score knowing that it could affect the school's overall rating. It also bothered her when she knew her peers had decided not to take the CSAP exams or when they took the tests but did not put any effort into it. She mentioned this irritation a few times during our interviews.

Insights/Unique Perspectives

Jackie had a problem with the CSAP exams – they were too broad in their scope. As a test designed to be administered to students with a large range of talents and abilities, the CSAP did not seem to her to reflect everyone's strengths. "I think you should give a test based on the people attending the school," she remarked. She felt the test should reflect the population at the school as well as what has been taught at the

school. Once again she was looking for a reason the test might give back something to her.

Conclusion

Other participants came to these interviews ready to offer up their discontent regarding the CSAP exams. Jackie came having thought little about her attitude regarding the tests. As she talked, she uncovered more of her frustrations and resentments. Her loyalty to school, her commitment to doing her best remained, but under this layer of commitment bubbled dissatisfaction that finally surfaced only after she had finished her final CSAP exam and had a chance to discuss her experience.

Vignettes

Outrage.

Jackie arrived early for the state writing exam; her pencils are neatly sharpened on the desk in front of her. After receiving positive results for several years she is certain she will do well today. Despite her confidence a slight amount of nervousness is evidenced by the way she fidgets with her pencil.

Just after the proctor begins delivering the testing instructions a handsome boy with long curly black hair glides in and takes the empty seat next to her. She knows who he is but she rarely interacts with him. He is popular among a group of her peers who are less academically focused than she. As he slouches in his chair he pushes his hair out of his eyes, apparently ignoring the proctor as she reads the testing directions.

When the proctor finishes and has given the go ahead to begin testing, the curly-haired boy quickly sets to work writing on the answer form. At first Jackie is impressed with the confidence he applies to the task. She soon realizes that he is just writing random words on the page, completely unconnected to the writing prompt. Jackie glances over as he catches the eye of one of his peers and smirks. Catching herself becoming agitated, she forces herself to suppress her anger and returns to her essay.

A few minutes later when she checks on him his head is on his desk. It is evident from his breathing that he has fallen asleep. She fights resentment again and returns to her work.

Narrative #5: Audrey

Initial impressions/ Description

Audrey, a pretty girl with long dark hair and eyes that were startlingly large, entered the counselor's office I was using for interviews, gave me a quick appraising glance and then slouched into a chair. I interpreted her body language to say, "OK, let's get this over with. Start with the questions." In the first few minutes of the interview I was not certain if she would reveal enough of her thoughts and opinions to me to make the interview useful. She warmed to the process and by the end of the first interview was beginning to open up a bit.

She did not show up for her second interview so I solicited the help of the secretary in the counseling office to help track her down. I found her relaxing with some classmates in an outdoor courtyard. She neither apologized nor complained as she trailed me back to the office for the interview. Surprisingly, she was more relaxed and forthright in this conversation. As the interview progressed she leaned forward, established eye

contact more frequently and began to really tell me about how she felt about school and testing.

Background/ Interests

The remote mountain community where Audrey lives is 40 miles away from Silverheels High School. While she appreciated some of what her unique living setting had offered her – natural beauty and a peaceful setting, she felt she has missed out on some things her peers in town had experienced. Audrey liked stimulation. She liked movement, music, conversation, and human interaction. She believed her mountain upbringing had denied her opportunities to engage in activities that would allow her these experiences. So while she recognized the benefits of living in the mountains, she also felt stifled by that environment.

Fortunately for Audrey, she excels in some areas. Movement is important to her. She described herself as a good athlete, a skillful basketball player and an enthusiastic dancer. When I spoke to her she was preparing for a dance performance. I asked her if she thought she was a good dancer. “Everybody says I am,” she replied. This contrasted with the negative feedback she frequently received from her testing experiences. There were times when she knew her dancing was not at the level she would like, but it still brought her joy. The movement itself was satisfying and the positive feedback she received felt good.

Testing History

Audrey was thoughtful, articulate and obviously quite bright, but school had not been easy for her. Reading, math and spelling had been elusive content areas for her. In her day-to-day schooling she had compensated for these deficiencies. She asked teachers

for help when she needed it. She had learned to access information sources or rely on peers when she needed some support. It was when she had to take a test that her learning challenges confronted her most dramatically. During these experiences her usual support strategies were removed, and she had to face difficult reading text or math equations on her own. She wrestled with the content, tried to make sense of it. Sometimes, the harder she examined a question the more obscure and inaccessible it became. “I know how to do it; it’s just finding the equation and the right way to do it. It just kind of mixes in and then is lost, so a lot of times I have to look things over and over and over or have an example in front of me,” She told me.

It isn’t surprising given Audrey’s testing history that she did not think CSAP results provided an accurate view of the person she was. She wanted to think she was smart and her parents had tried to convey to her that she was intelligent, but when test scores came back year after year with low scores it was difficult to reconcile those with these positive messages. She wondered what happened to her knowledge when she sat down to take an exam. She realized that what she knew was much greater than she was able to communicate through answering test questions. Language got in her way. She struggled with making sense of it and she was constrained when she had to respond in narrow and technical terms. She had an apt metaphor for her frustration: “It’s lights out and you can’t find it, it’s like you’re running for the answers but you can’t grasp them.”

It has been tough for her to keep getting negative messages from test results. She told me, “It’s just like now for all the tests that we do poorly on, it’s like, oh well, I guess you don’t know the material or anything....Oh, we don’t know it so we’re dumb.” She steeled herself against the pain by putting up a “shield” to protect her. She pretended the

persistent test feedback did not hurt, that she did not care. Discussing her less test adverse peers she commented, “How do they know it and I don’t?” This question appears to have haunted her for a long time.

Audrey was lucky to have parents who understood her educational challenges. She said she had been able to talk to them freely about her learning frustrations. They had been successful in convincing her that she should go to college. Despite her educational challenges, she said she was “definitely” going to go to college. Understanding as they may have been regarding her educational struggles, Audrey and her parents had disagreed on the importance of school attendance. She said her dad had often told her that he only missed a single day during his high school years. Audrey held a more relaxed view about attendance, staying away from school when she had “something more important to do.” It was not surprising she was sometimes not eager to attend school given the messages she received there were not always positive.

Testing Reflections

Audrey’s views about testing were similar to those of other participants. She felt the test did an inadequate job of assessing her knowledge and that it did not allow what she knew to be displayed. She also resented that the test was imposed upon her and her peers when it had no bearing, either positive or negative, on her future. When I asked her what she would tell state legislators about the exam, she said, “I would tell them that CSAP testing is unnecessary in schools because kids don’t care about it, they don’t try. If kids don’t care about something then they’re not going to try hard and the results won’t be necessarily factual, they’re not going to be true.” This struck me as one of the more astute observations made by my participants.

Audrey's opinions about her relationship to testing shifted back and forth during our discussions. At times she acknowledged the pain that testing had caused her. Lowering her shield, she told me about how much it hurt to fail. At other times she struck out at the test and blamed the shortcomings of the test for her failure. As she moved back and forth between acknowledging the pain and railing at the test she also shifted between stating she did not care about the test and did not try, to admitting that she really did try but still received unsatisfactory results. When I suggested the latter might be closer to the truth, she responded, "Yeah you are right. It's just, yeah, it gets so frustrating and just difficult, so... Like the more you try, the harder it gets sometimes, so you just give up."

Peer Responses

When Audrey was younger she only had one or two close friends. Living in the mountains and attending school in town made social encounters more difficult to arrange. That may be why as a high school sophomore friends and group belonging were quite important to her. She told me that her friends had high energy and were loud and independent. Among this group of friends she said she was the highest achiever. She thought she was the one who was most likely to try hard in school and she was one of the students in this group who received the highest grades. According to Audrey their attitude regarding CSAPs and school in general was less accepting than hers. She said many of her friends did not try at all to do well on the CSAP exam. They merely filled in random responses. "They don't care," she said. I think it may be a testimony to her parents' attitudes about education that she retained her comparatively positive attitude about school in the face of her peers' disdain.

Insights/Unique Perspectives

Audrey's strong link to peers was mirrored in her connection with other people in her life. She felt she could talk openly to her parents. She also said she liked to have good relationships with her teachers. She had a need to connect with them, to know them as people, to go beyond just talking to them about course content. It seems to me that she needed this connection to help her bridge her struggles with content. She demonstrated this need for connection during our interviews. She was the only one of the participants who asked me about my stance on the CSAP exams. During our first interview I demurred, but when she asked again at the end of the second interview I let her know my views.

Relationships, she has learned, not only fed her soul, they helped her learn more easily. She flourished in learning settings where she could work with other students. She recognized that individual strengths and understandings of group members added up to something greater than any of its members had individually. Audrey described it this way:

My mind definitely works a lot easier when there's more than one person working with me and it just seems to get a lot more. You can expand your knowledge past theirs and then they expand past yours and it just kind of keeps growing and growing and growing.

These peer connections sustained Audrey and helped her get through school but in testing situations she still had to fend for herself.

Living daily with silence as she had in the mountains, I expected Audrey to have a different relationship to silence than her peers. I was surprised that she, like my other

participants, was not at home with silence. She rarely sought it out and when she encountered it, it was not comfortable. She either had music or television on when she was at home and she confessed to me that she routinely fell asleep to the sound of the television in her room. If she was uncomfortable with silence at home, she was even more uncomfortable with it in testing situations. Sitting in a room full of students, and forced to sit in imposed silence, she said she wished she could break through it with a disruptive word or action.

Conclusion

Audrey received academic support from her parents and she got some support at school for her learning challenges. Despite this support, her learning at school was not easy. For many of her peers the CSAP was just another adult contrived annoyance. For Audrey though it went beyond that. Confronting the tests, stripped of her support mechanisms, she was up against what she conceived to be obscure and disconnected material. The exam reminded her of how hard learning could be. I share her relief that at least this source of pain is over, I admire her strength to continue facing the challenges she has yet to face.

Vignettes

Loss of an ally.

Audrey's brother had Mr. Bunting when he was a sophomore. Her brother shares her academic struggles. He has trouble with reading, math and spelling in the same way she does. When her brother recommended a math teacher she knew that teacher would be a good fit for her as well. And it wasn't just her brother who said Mr. Bunting was a great

math teacher. Lots of her peers said he was cool. She didn't know what her first day at school would bring, but at least she knew this one class would work out for her. He was a teacher who had a reputation for helping kids understand hard math concepts.

At least that was the way it was supposed to be. The day before school was to begin Mr. Bunting was killed in a tragic car crash. Those students who knew him were shocked and sad; many openly cried. She didn't know him so she didn't cry, but she mourned the loss of having him on her side in her fight against difficult math curriculum.

Maybe nobody could take Mr. Bunting's place, but the substitute who replaced him sure did not fill his big shoes, at least not in her mind. When she asked him for help before their first big test, he brushed her off. She was disappointed. Her friend felt the same way about him. To protest his reluctance to help them, they both turned in their tests with their names on the top but with none of the responses filled in.

Narrative #6: Maria

Initial impressions/ Description

Maria seemed shy as I greeted her and explained the interview process. Her brown eyes conveyed absolute kindness. As I began asking her questions I appreciated the care with which she responded. Because she was so sweet and honest her answers were heartfelt and genuine. She told me how hard testing had been for her and how hard she needed to work to do well in school. Her struggles had not made her bitter. She seemed willing to accept difficulties buoyed by emotional support from her family and friends.

Background/ Interests

Because Maria had lived in numerous places in the Denver and Ute Creek area she had been required to make many transitions. She had attended six elementary schools and two junior high schools. Even before she began school she had to adapt to a number of daycare changes. Some kids take frequent moves in stride, for Maria they were not easy. She described to me the anxiety she felt when she had to adjust to a new day care or school setting. She discussed how nervous she felt when she met new classmates and teachers. It required a great deal of emotional capital to make these new acquaintances all the while knowing that she would likely have to let them go when she made her next move. Happily, at the time of our interviews she said her family was through moving because they had recently purchased a home in a small town north of Ute Creek and planned to stay put for a while.

It is not surprising; given Maria's frequent moves, that she derived a great deal of support from her family. She told me that her family had fun together. They liked being around each other and they shared a supportive relationship, frequently making each other laugh. She has a younger brother and sister, both of whom Maria described as "high achievers." Maria, in contrast to her siblings, had to work harder to succeed in school. She recognized that their path through school was easier than her own. Fortunately for her, her parents recognized their children had different strengths. They conveyed to Maria they were content with her test results and grades even though they were not as high as those her siblings brought home.

Almost as important to Maria as her family were her friends. When describing her group of friends she mentioned they were funny and that they had good values. She liked

that they were able to make her laugh. It was of some comfort to her that her parents approved of them. They had even remarked to her that she had great friends.

Her relationships with peers had not always been a source of comfort, however. When she was beginning elementary school she struggled with peer relationships. She remembered being teased because she had a scar on her hand. Adding to this, her frequent moves made it difficult to form long-lasting relationships with classmates. Given this history it is not surprising that her present peer group was so important to her.

Testing history

Math problems were difficult for Maria and she described herself as a slow reader. These factors were compounded by test anxiety. Before a test she had trouble eating and her stomach remained unsettled until she took the test. She told me that the anxious stomach subsided when she was taking a test, but this was replaced by a more vexing problem. Her brain would freeze. She could not remember equations and she had trouble processing the test content. Her anxiety interfered with her ability to process information that would normally not be difficult. Despite her anxiety and her perceived poor performance Maria still generally received scores of “proficient” on the state tests. This leads me to believe that she was a better student than she thinks.

Tests make Maria nervous but so do other new experiences. Just as she felt anxious in new daycare settings or in new schools, she had trouble with tests when she did not know what the content was likely to be. She told me that if she knew what was to be on a test she was much calmer about taking it. Even though she said she disliked reviewing for tests, she knew if a teacher prepared her for the content she would be less stressed and would perform much better on the test. Because it was difficult to review for

content on the CSAP exams they were much more stressful for her than the tests she took in her classes.

Testing Reflections

Despite her test anxiety Maria did not resent the CSAP exams. She accepted them, viewing them as just another obstacle to overcome. “I think the test makers had good intentions,” she told me. Likewise she did not think that the test designers were trying to trick her into supplying the wrong answers to test items. She did not believe the instruments should be a one-size-fits-all situation, however. She thought different students with different backgrounds or learning styles should be able to take different versions of the test.

Peer Responses

Maria did not say much about the responses her peers had to the CSAP other than discussing how they felt about the testing environment. She thought most students were uncomfortable in the formal and carefully controlled test setting and she thought her peers would enjoy being able to listen to music while testing.

Insights/Unique Perspectives

Maria had trouble with the restrictions imposed by the test. Like other participants she was uncomfortable with the climate in the test room. The testing room was full of kids who did not want to be there she pointed out. Many were nervous and most were not comfortable sitting in imposed silence. “It seems that people play off of each other’s emotions,” she observed describing the testing situations. She believed the tension in the

room was the result of the cumulative anxiety and discomfort of the students assembled there.

Conclusion

It is easy to see why Maria has attracted a cadre of friends. She is sensitive and sweet and she seems to be free of anger regarding her schooling or testing. She could have lashed out against the CSAP exams. They do cause her physical and emotional distress and they provide her with an annual reminder of how poorly she performs in testing situations. Despite these things, she did not express any hostility towards either the test or the test makers. She agreed to talk to me because she wanted people to understand how students feel about the test, “for not just me but a lot of people.” She wanted to speak not just for herself but also for her generation – her peers and classmates. She wanted to get the message out but without any accompanying hostility. She did not say so directly, but she implied that she hoped people would listen to what the voices of youth had to say about this test that had caused her such distress.

Vignettes

Self-talk.

“You may begin,” states the test proctor. Maria opens her CSAP booklet to the correct section. She begins reading the first piece, a non-fiction selection. Her eyes scan the words but nothing registers. She looks around the room, it is so quiet, unnaturally quiet she thinks. She is aware of tightness in her shoulders and arms. Her stomach is more settled now than it was before the test but it still isn’t right. “ Maybe I’m hungry,” she thinks. She tried, but could not eat any breakfast.

A recalibration.

“The words, I need to focus on the words” she reminds herself but then stops mid paragraph and turns the pages, counts them, determines the number of responses she needs to complete. Back to the paragraph, she breathes in twice and looks quickly at the clock. “I can do this,” she thinks. “I always get through these things. My scores aren’t great but they are OK. I can do this.” She breathes in again and starts reading the paragraph again, slowly. The words start to make sense. “This isn’t so hard,” she thinks as she bubbles in a response to the first question.

Themes

In the following section I present seven themes that emerged during the analysis process. I describe the themes and discuss how each is related to responses from individual participants. I have used frequent participant quotes to illustrate this relationship.

School Loyalty and Test Compliance

“I don’t think they really understand that if they do go to this school they should show how intelligent they are, and by just kind of blowing it off they’re not doing that at all.”
— Jackie

Early in the interview process I encountered a division among my participants. Some stated that they were motivated to do well on the CSAP exams at least in part because they wanted to represent their school in a positive light. They wanted their teachers and school to look good. One participant, Jackie, had particularly strong feelings about this. She was quite contemptuous of students who did not share her sense of loyalty to Silverheels High School. She complained:

I kind of think it's unfair to the school. Here they are attending school that they're getting funding for and if they want their school – if they want to go to a good school then they should do good on their exams and show what they really know.

Dennis and Audrey did not share Jackie's conviction. Dennis said that although he felt a connection with his teachers and to his IB peers, he did not feel compelled to perform well on the CSAP exams for the sake of the school. Likewise, Audrey, when asked if she felt she should try her best for her school, stated she did not harbor any such loyalties. In fact, she seemed to revel in her contrary stance. She did not feel she owed the school anything.

Audrey had struggled with school tasks, and particularly with testing, throughout her school career. Reading, mathematics and writing conventions had eluded her as she moved through school. So it is not surprising that she responded with anger when I asked her whether she felt she needed to perform well on the CSAP exams for the sake of her school. "I mean I don't really care about how our school is ranked. I'm in the school yes, but my personal grades don't have anything to do with how the school is ranked in the district," she remarked.

In contrast, Maria, another participant who had learning difficulties, was more charitable in her attitude toward school. She admitted that she had struggled throughout her education and did not do well on tests but there was no anger in her tone, no resentment. She wanted to do her best and would have been happy to perform in a way that would have elevated her school's standing, but her testing capabilities were not strong enough to do so.

While Maria and Audrey struggled with the CSAP exams, the other four participants scored well. One of these participants, Dennis, posted excellent scores on his exams despite denying any loyalty to his school. He said he put forth good testing effort in order to maintain his feelings of self-worth.

My participants varied greatly in their attitudes about performing well on the CSAP exam for the sake of their school. Jackie was fiercely loyal to Silverheels High School, Dennis claimed no loyalty to the school, Audrey bridled at the notion of loyalty, Thomas, and Michael were happy to contribute their test performances to help the school's rating, and Maria would have gladly joined had her scores been sufficient to help out.

Test as a Measure of the Student

"I would say I got all my bubbles and some answers and this is supposed to be a reflection on me and how I can learn, and I don't think it is."

— Dennis

One of the questions I asked the participants was, "What does the CSAP say about who you are?" No one thought the test provided an accurate picture of him or her as a student. Michael replied that the test results merely revealed, "What you remember." Jackie commented that the test was merely a measure of motivation. Similarly, Maria thought the CSAP results showed that, "I probably don't try as hard as I could or should." Dennis thought the exam inaccurately measured his math ability and failed to correctly gauge his writing ability because the prompts were "unrealistic."

The CSAP exam contains short and extended response items. It is not an instrument that relies solely upon multiple-choice responses. As such, the exam relies

upon the writing skills of the test-taker to communicate the depth of their understanding. Facile writers like Thomas and Jackie can easily convey their knowledge to the test readers. For the other four varying levels of writing competence and writing confidence interfered with this process. Michael balked at expressing thoughts and feelings he regarded as too personal. Dennis confessed that he was a formulaic and restrained writer. He was quite aware his level of expression was lacking. Maria labored through reading questions and then wrote anxiously, never at ease in the test setting. Audrey liked to write but knew that her poor spelling and high number of grammatical and punctuation errors deterred from her score. Although the reasons varied, these four students all felt their writing interfered with communicating the full extent of their knowledge. Dennis went beyond this assessment and suggested that the test was ineffective because it was somewhat contrived and artificial. No participant felt the test revealed much about what they knew or could do.

Personal Relevancy of Test

“...CSAP is kind of pointless to me just because the stuff that you learn in school and outside of school, you can't put in a test.”

— Audrey

All my participants except Thomas voiced a critical opinion regarding the personal relevancy of CSAP testing. Some of this criticism appeared to be well practiced, as if they had participated in numerous discussions with peers and family about the personal importance of testing. At other times their critiques were more spontaneous, and sounded as though they were being formulated as they spoke with me. Some did in fact mention that they had not thought much about their testing sentiments until they were interviewed.

Every participant but Thomas commented that the CSAP exam was unrelated to his or her future. For the most part these students had difficulty linking their performance on the test with any aspect of their lives. They could not see how their future was related to CSAP scores. In contrast, they could point out how the SAT, the ACT and their cumulative grade point average could be directly connected to their acceptance to a college or university. For these kids who were looking ahead at their own future, the CSAP exams exerted precious little sway.

Jackie and Dennis mentioned other problems with the relevancy of the test. These two were bothered that their test results provided them with little specific feedback about their performance. They received a ranking but nothing that would help them improve their performance on future tests. “All I receive is this little list of numbers which means nothing to me I think if anything the test should teach you something,” commented Dennis. When one of their high school teachers returned a test, these kids received useful information about where they excelled and where they fell short. The CSAP results did not give them the same sort of data.

To say these students welcomed tests from teachers or college entrance exams would be an overstatement, but they took these tests without resentment because they understood their importance. They could not make the same sort of link between the CSAP and their education or with their future.

Motivation

“...The difference is I tend to care a lot more than of the other classmates. Like I said, I try hard no matter what.”

— Michael

Of the six students I interviewed five expressed they had taken CSAP exams with the intent of doing well. These kids were clearly motivated to perform well on the test. One of these students was Jackie. When asked about what her test scores said about her, she replied, “I think it kind of shows that I care probably, or that I am motivated to do it...” That was certainly the overall message she conveyed to me. Twice during the interviews, however, she admitted that she struggled with a rationale for doing well on the CSAP exams. “I kind of feel like it’s annoying, a little bit, because it doesn’t really – there’s no consequences for you to do bad and there’s no reward if you do good.” Later she revealed her discontent with her preparation for the exams, “Right, and the staff says, ‘Oh, this will be on the CSAP, you need to know how to do this.’ And we’re like, why? If it’s on the CSAP, we don’t care about the CSAP.” Jackie worked hard and performed well on the tests but it wasn’t because of any external reward or because she held the test in high esteem. Despite her stated ambiguity – a conflict between her ingrained desire to do well and her critical views about the exams, motivation appears to have triumphed.

Thomas, Dennis, Michael and Maria described testing attitudes that displayed to me that they were motivated to do well on the exams. None of these students communicated that they had the sort of reservations about performing well that Jackie had described. Thomas, Dennis and Michael also expressed confidence in their test-taking abilities. Maria stood apart from the other three because even though she wanted to do well, her test taking abilities interfered with her performance.

The participant whose attitude about the exams was clearly sour was Audrey. Some participants had mentioned that many of their peers were not motivated to take the exams and that they had poor attitudes about them. Audrey’s motivation seemed to mirror

that of these other students. According to the students, these peers either did not take the test at all or they did not take it seriously. In the words of Audrey, "...they'll do it. They just don't try, they'll just fill in the bubbles, they won't even read it sometimes." I had the impression when she said this that she was talking about herself as well as about her peers. Speaking more directly about her testing attitude Audrey said, "It's not a happy, happy thing, it's not a joyful thing."

Despite telling many tales about unmotivated peers, all but one of students in this small group described themselves in terms suggesting they were motivated to do well on their CSAP exams.

Test Anxiety

"The CSAPs I've taken these last three years have been a breeze."

— Thomas

My participants were divided on the issue of test anxiety. Test anxiety was not a stated concern for Thomas or Jackie; Dennis, Michael, Audrey and Maria however, reported experiencing various degrees of anxiety while taking CSAP exams.

Michael described the mildest of these cases. He struggled with writing and disliked testing responses that required him to write. Dennis also had trouble with writing tasks. He admitted that writing and spelling had always been difficult for him and that these limitations often caused him to be disappointed with his finished work. For these two boys test anxiety seemed to be restricted to areas of the test where they had to write extended responses. In other testing areas, they were confident in their abilities.

Audrey was another student who admitted to suffering from test anxiety but this admission came after some initial bluster. In my first interview with Audrey she

complained about the testing process and other aspects of her schooling. Her tone was somewhat defiant, perhaps even angry. It appeared to me that she had constructed a tough shell to protect herself from the negative messages and pain that she too often experienced when testing. During our second conversation, perhaps because she felt more secure, she allowed that testing for her was not a happy experience. “You get frustrated because you want to know it, but you can’t. So you try to do a problem over and over and you can’t figure it out. And it’s just – frustrating and then you want to give up.”

Unlike Audrey, Maria made no attempt to conceal her anxiety regarding testing. She was not angry or resentful about being tested. For her testing was just something she had to endure. Obediently and without any animosity she participated in the painful process year after year. “You’re so stressed that you’re just like – you feel that you know it all and then you get to the test and it’s like you’re completely blank, like you don’t know.”

While Audrey and Maria clearly experienced anxiety, Thomas and Dennis experienced some mild distress when they took the CSAP math exam. As advanced math students they were two or more years ahead of their peers. This meant that they had forgotten some of the algorithms and formulas that were fresh in the minds of their peers. This said, I did not get the impression from talking to these boys that their test score had suffered very much because of their lack of recent exposure to content. It is likely they more than made up for this deficiency by getting more difficult questions correct. For example, Thomas explained to me how he had used calculus to solve a word problem asking him to determine fence post quantity. Another thing that bothered both boys was

they had felt unprepared to answer some astronomy questions that had been on their eighth grade science CSAPs. These two, used to success, were quite uncomfortable when they were presented with content with which they were unfamiliar.

My participants exhibited a wide range of emotional responses when confronted with the CSAP exams. These ranged from the mild discomfort described by Thomas and Dennis when they couldn't remember a math algorithm to the strong anxiety that caused Audrey and Maria to feel their brains had shut down when they read difficult test questions.

Human Connections and the Unknowable They

"Like the way they word the questions, I guess it throws me off sometimes."

— Maria

After reading and rereading the interview transcripts I was struck by how distant and anonymous those responsible for imposing the CSAP exams were to these students. Some students understood that state legislators were responsible for imposing the testing, and some knew that the State of Colorado Department of Education employed a cadre of readers to grade the exams, but for the most part the people responsible for high stakes tests were remote and unknown to them. Nameless test makers constructed questions and another set of unseen people graded their exams. Invisible and unknown, the students could only identify them as "they." *They* passed laws mandating the test, *they* made the test, *they* read the essays, and *they* give the scores.

Thomas, who was likely the most enthusiastic test-taker in my group of participants, attempted to reach out across the void that separated him from these unknowable individuals. One of the prompts on his CSAP writing test asked him to write

a letter to someone he would like to talk to. Thomas explained his response, “I wrote my letter to the CSAP grader to ask him questions about, you know, why he got chosen for the job and why he took it.”

Unfortunately for the students, *they* do not write back. The unknowable they provide no reason for students to perform at their highest level. Michael remarked concerning the test-makers, “Nobody ever came out and said, “This is why we’re testing, it’s important.” He wanted those responsible for instituting the CSAP exams to step forward and explain themselves.

In contrast to the “unknowable they” are the people in these students’ lives who help them in various ways in the face of high-stakes testing. Teachers provide them with the background they need to tackle the tests and give them a reason to perform. Jackie mentioned that she tried hard on tests so that her teachers would think well of her. Parents provided support and a sounding board for student concerns. Parents also pushed the students to do well in school and insisted that their children take the exams. Their peers shared the experience with them and through discussion, helped shape the participant’s attitudes about the testing process. It is in the face of these rich supportive interactions that the mute and unknown responsible for developing and scoring their tests appear in such stark contrast.

Relationship to Silence

“It’s such a low energy room and you need the energy around you to learn.”
— Audrey

This finding was the most unexpected of the thematic elements I encountered. Early in the interview process two participants mentioned being uncomfortable with the

enforced and artificial silence they encountered during the testing process. Their comments led me to ask other participants to explore this topic with me during the second set of interviews. I had already conducted two interviews with Thomas when I made the decision to discuss silence with the participants so he did not have an opportunity to respond to this topic.

Dennis, Michael, Audrey and Maria said that silence during CSAP testing was uncomfortable for them and said they avoided silence at home and during other times in their lives. To prevent silence from intruding they turned on television, engaged in conversation, or listened to music. Their lives were filled with sound nearly every waking minute. Just as primitive humans sought fire to keep the beasts of the night at bay, these students employed iPods to ward off silence in their lives.

Jackie stood apart because she was more ambivalent about silence. She acknowledged that her peers had difficulty with silence during tests but she said that she sought silence in her own life, especially when she was reading. Her responses indicated that silence during the CSAP testing was not a hardship for her.

These kids go through their lives with almost constant auditory stimulation and then they are put into testing situations where they must sit for prolonged periods with very little sound stimuli. Dennis remarked, "Silence is so loud." Maria remarked, "It's kind of hard for kids to sit there and concentrate and keep it in focus – like for me, I learn better when I have like music or something in the background – or even just let people talk."

Individual Participants and Thematic Elements

In this section I revisit each of the participants to summarize how the themes in the previous section relate to them as individuals. In this way I can connect the themes back to the participants.

Table I provides similar information. The reader can refer to this table to quickly connect individual participants with a theme or those with a quantitative urge may count how many participants contributed remarks in support of the various thematic elements.

Table I

Thematic Elements Discussed or Exhibited by Participants

Theme	Thomas	Dennis	Michael	Jackie	Audrey	Maria
Described feelings of loyalty to school	X		X	X		X
Said CSAP does not describe them as a student	X	X	X	X		X
Indicated he/she was motivated to do well on CSAP	X	X	X	X		X
Contested personal relevance of CSAPs		X	X	X	X	
Described some level of test anxiety		X	X		X	X
Anonymity of test designers and graders discussed	X	X	X	X	X	
Expressed discomfort with silence		X	X		X	X

Thomas

Thomas is the type of student test designers dream about. He was internally motivated to do well on the CSAP exams because of his loyalty to school and because of his inner drive. This motivation was magnified because of the strong pro-education

messages he had received from his parents. He did not question the relevance of the CSAP exams. To him it was just something he had to do. Because he was confident of his abilities and because he had a history of doing well on tests, he did not suffer any anxiety while taking the CSAP exams.

Family, teachers and peers are important to him. He emphasized in our discussions that he had received frequent messages from his family about the importance of school. Despite the importance of these human connections, he did not express any animosity regarding an absence of a relationship with those individuals who constructed and graded the test.

I have shown in Table 1 that Thomas did not indicate that silence was uncomfortable for him. This is somewhat misleading because he was the only participant that I did not ask about this relationship.

Dennis

Like Thomas, Dennis was compliant and a good test taker but his background made him a bit more critical. It was not until his high school years that he found a peer group with which he felt close. This lack of connection within schools and his frequent moves may have contributed to his ambivalence regarding school loyalty.

As a highly analytical person he turned a skeptical eye on the testing process. He questioned the relevance of the test and he was clearly critical of the importance of the test to his future. Of all of the participants, he was one of the most vocal critics of the process that he felt isolated him from a connection with graders and test makers.

Anxiety is probably too strong a word to use to describe Dennis's mild distress over his less than stellar spelling and writing scores. These deficiencies did, however,

cause him some trepidation when he faced CSAP questions that required extended writing responses.

Dennis admitted that silence during testing could sometimes be distracting but he also thought it helped him to focus. He believed eliminating the requirement for silence during the exams would, “undermine a little of the authority.”

Michael

Michael is loyal and motivated. He told me he wanted to do well on the CSAP exam because he did not want to see his school taken over by the state and turned into a charter school. His level of motivation was evidenced by his professed ethic, “If you are going to do something, do your best.” These positive attributes did not stop Michael from questioning the validity and relevance of the CSAP exams. He, like Dennis and Jackie, was skeptical about whether the state exams revealed a true picture of his learning. He was particularly critical of writing prompts that he felt were artificial, and in his words, “... not very thoughtful.”

Michael, more than any other participant, asked direct questions about whether the test was relevant. He wished that those responsible for the testing process would tell him why it was important for him to invest his time and effort into preparing and taking the exams. He also wanted to know why the testing was worth losing two weeks of instruction

Michael was not a strong writer. “Whenever there’s writing on anything, I kind of moan inside.” Because of his struggles with the written word, I classified him as having some test anxiety, although I must remark, this mild anxiety hardly compares to that experienced by Maria or Audrey.

Michael told me that he liked to turn up the music loud in his room when he was doing his homework. He turned it up loud enough so that it “drowns out everything else.” He thought everyone would enjoy an opportunity to listen to music during the exams. “I’ve got an idea,” commented Michael, “What we should do is they should organize classrooms by who likes what music.” Michael would choose the room playing classic rock.

Jackie

Jackie is the portrait of a model student. She excels in school, she was involved in student government and other extra-curricular activities and she was scornful of students who did not take the CSAP exams or did so without exerting any effort. Given these sentiments, it is somewhat surprising then that she was also a sharp critic of the testing process. She directed her keen intellect towards questioning the relevance of the mandated testing. She decried the instructional time wasted to prepare for the tests as well as a system that failed to provide her with direct feedback regarding her performance. Despite her harsh views regarding the testing process, she was extremely loyal to her school and performed at the highest level on the CSAP exams. She was confident and capable and did not reveal any academic weaknesses to me during the interview process, so it not a surprise that she did not mention experiencing any anxiety during CSAP testing. Likewise, Jackie did not mention having a problem with silence during testing. She did recognize that this was a difficult for some of her peers.

Audrey

Loyalty to school provided no reason for Audrey to perform well on the CSAP exams. In her mind the test was an objectionable and invalid task that was merely to be

endured. She felt no obligation or motivation to persevere through difficult portions of the exam. The best thing to do in her mind was to just finish and get it behind her. She talked tough early in the interviews but then admitted that the test was mentally stressful for her, perhaps because it reminded her of her previous difficulties with academic work.

I think that Audrey was the participant who most actively sought connection with others. She spoke about how jarring the death of one of her teachers was to her, she talked about the importance of her peer group at school, and she asked pointed questions of me. She wanted to know what I thought and where I stood on issues of importance to her. Audrey said her performance might not improve if she had a relationship with the test maker, but her attitude would likely be better.

It was when interviewing Audrey that the topic of silence first came up. I was surprised that this girl who lived in a quiet rural setting expressed the strongest attitudes about having a discomfort with silence.

Maria

Maria shared Audrey's difficulty with academics and tests but she did not display any hostility towards the test. She was willing to do her best as far as her abilities allowed. She tried to perform on the tests but sometimes her mind just went blank. This may be related to her learning disability or it may be because of the test anxiety she admitted to suffering when testing. Her mental lapses during testing made her wonder whether the test was a valid measure of her abilities.

The personal relevance of the test is not something that Maria addressed. For her the test was something that she anticipated with trepidation, endured, and then put behind her until the next year. Since Maria did not contemplate the nature of the test or her

relationship to it, she also did not discuss having any difficulty with distant and unknown test makers and givers.

Maria is from a large family, a noisy group generally, according to her. She attributed this frequent exposure to plenty of sound to her discomfort with the silence she encountered while taking the CSAP exams.

Summary

In this chapter I have presented individual narratives followed by a description of seven thematic elements. It is the nature qualitative research that these stories and themes are my own construction, therefore, to add a dimension of trustworthiness to the study, all the participants were given an opportunity to comment upon the narratives. Three participants responded. One thought his was, “a bit dramatic but got the point across.” He went on to suggest that I change a word in the narrative that “was a tad unprofessional.” He concluded by congratulating me, “Overall, nice work” and wished me good luck. The other two had less to say; mainly communicating they were pleased with the narratives and did not see anything they wished to see changed. The affirmation of these three participants, while comforting, still does not mean that my portrayals are accurate portrayals of the individuals, just that they are close enough that these three participants could sign-off on them. As for the three who did not respond, I will never know what they thought of their narratives. I would like to think that they would have spoken up if they felt they were far off the mark.

I can make similar remarks regarding the thematic elements. These elements are clear and distinct to me. After repeated reading and rereading these are the themes that remained in my mind and appeared and reappeared in my notes. Now they have become

the dominant component of my investigation. A different researcher bringing a different view would likely have found other themes common to the same set of participants.

These, however, are the themes I have found and have defended and these are the themes I will further explore in the chapter five.

CHAPTER 5: DISCUSSION

Introduction

This chapter is divided into two parts. In the first section I revisit seven thematic elements I identified in chapter four and link them to a central metaphor. I also discuss implications the elements have regarding educational policy and explore possible topics for future research. Section two is presented as a panel discussion with four participants representing four stakeholder groups. The panelists address six questions regarding high-stakes testing. The purpose of using the panel format is to place the findings of this study into a broader arena where they can be challenged by contrasting views. The content of the panelists' narratives was influenced by data I collected for this study, a review of the literature, and personal experience.

Discussion of Thematic Elements

In chapter three I described the legislative process that birthed No Child Left Behind. The process was largely top-down and directive rather than collaborative since the legislators solicited little input from educators or academics. I characterized the role of legislators in this effort as paternal. Their approach seems to convey the message that they knew better than educators what was good for the nation's schools. As I pointed out in my literature review, the directives were first passed from the Federal government to states, then to districts, next to schools and teachers and then finally to the students who sit for the required exams. Given this chain of directives it not surprising that students

respond to high-stakes testing in similar ways as they respond to directives from their parents. Because of this similarity I think it is appropriate and instructive to use a parental metaphor to explore the thematic elements identified in the last chapter. Employing this device I will introduce each discussion of the implications of the thematic elements by linking the particular element to the parent-child metaphor.

School Loyalty and Test Compliance

Some children are fiercely loyal to their parents and others seem to be born to challenge their parents at every turn. In a similar way the participants in my small-scale study expressed differing degrees of desire to perform well on the CSAP exams for the sake of their school. Their loyalty seemed unrelated to reported test performance. Dennis claimed to have little loyalty to his school yet he performed well on tests. Audrey expressed similar sentiments but did not do well on the tests. Loyalty to school may be a factor in test performance but nothing from my study points in this direction. It appears that other factors such as ability, peer influence, and motivation are more important to test performance than how students feel about their school.

Given that my study found no connection between loyalty and student performance I will not suggest any policy changes related to this element. It may be instructive for educators and administrators; however, to consider that any attempts to increase school loyalty for the purposes of increasing high-stakes test scores may prove to be fruitless.

Since this element did not appear to be linked to test performance I would not be inclined to investigate it further. Other thematic elements discussed here pose more compelling questions for future research.

Test as a Measure of the Student

A common lament among adolescents is, “My parents just don’t understand me.” Focused upon peer relationships and the pressures of maintaining those relationships, teens begin to look outside their families for behavioral cues and affirmation. At the same time they are looking outside their family for guidance, young adults are going through remarkable physiological and psychological changes that amplify their feelings of difference from their parents. In a similar way, students in this study questioned their relationship to high-stakes tests. They challenged the notion that the exams had anything to say about who they were as individuals.

The complex relationships between parent and child are difficult to separate into neat categories. The same is true of this thematic element. It is difficult to separate it from personal relevancy and motivation. If students felt high-stakes tests had personal importance they might be more highly motivated to perform well. In turn, if they felt the test was relevant, and if this improved their performance, they might also believe the test said something about what kind of students they were. That is a lot of ifs. The students I interviewed did not find the state exams to be relevant to their lives and they were not all motivated to do well on them. Generally they found the tests to be superficial and not a good measure of who they were as students. When a student, particularly an expressive student like Jackie or Audrey, takes a high-stakes exam she spends a great amount of time filling in bubbles. She fills out some short response items and completes some longer essay items. Several months later she receives her results in the form of a rating: advanced, proficient, partially proficient or unsatisfactory. It is understandable that she would be frustrated by the process given that she has difficulty discerning the purpose for

the tests initially and then when she finally receives an assessment of her performance it is reduced to a one or two-word rating.

Some of my participants said that when they were younger they had taken the exams uncritically but as their intellectual skills increased they had begun to question what the tests said about them as individuals. They had come to realize that they had knowledge and abilities the test was incapable of measuring. Viewed through the lens of Howard Gardner's theory of multiple intelligences (Smith, 2002) these tests appear to be skewed towards measuring linguistic and mathematical-logical intelligences while ignoring the artistic forms of intelligence. Dennis and Michael may have had creative insights regarding topics posed on the CSAP exams but they were unable to convey them fully because of their writing limitations, or, given the multiple intelligence view, because they were not as strong in the linguistic intelligence domain. In a similar way Audrey had talents, or perhaps artistic intelligence that was not illuminated by the testing process.

Two of my participants, apparently seeking more interaction with the test instrument, said they took liberties with their writing prompt responses by writing playful essays. I think they were seeking a deeper connection with those behind the instrument. Their actions allowed them to reveal more about themselves than the test required. If we again view student behavior through the lens of Howard Gardner's multiple intelligences it appears these students were pushing to expose untested intelligences to the test givers. Perhaps if the instrument were designed to be more interactive, and if it allowed students to reveal other talents, they would feel it said more about who they were as individuals.

Learning is complex and elusive. The high-stakes test designers have created instruments that provide easily quantifiable results but I think the resulting data only goes part way towards describing the learning required of today's students. I would like to see a school district or state pilot an assessment process that would assess life-long learning skills. Such a pilot would allow researchers to appraise if a shift towards assessing life skills would improve student attitudes regarding the tests. I believe students would be challenged by such an instrument, likely more so than they are by current instruments, and I think it would reveal more about the qualities of the students and thus more about what they have learned that is of value to them and society. If these changes were made to the test instruments "high-stakes" would take on a whole new meaning. The stakes involved would be high not because the futures of schools depended on them but because the skill sets tested would be so important to the lifelong learning of the students.

Personal Relevancy of Test

After children learn to walk, and as they learn to communicate, they realize they exist as beings separate from their parents. With this recognition comes the need to establish their independent identity. Two-year olds are notorious for exploring their newfound independence. As children acquire knowledge, judgment and experience they seek to obtain increasing levels of independence. With independence comes a quest for meaning. Some parents welcome this yearning for relevance while others view it as a threat to the values they have attempted to instill in their children.

The quest for relevancy was certainly evident among the participants in my study. Testing relevancy emerged as a theme long before I began reading and rereading transcripts. It was a recurrent chorus I heard distinctly during the interview process. This

was not a nuanced message requiring repeated and careful reading of the data. Four of the participants directly addressed their issues with state testing. They stated they had difficulty seeing a connection between the exams and their lives.

This desire for meaning in schooling is explored by Unger (2003), a study I mentioned in chapter three. His findings suggest that students desire to express opinions, ideas and beliefs as part of their schooling process. He found they also wanted to have a voice about the direction their education was taking. Unger, however, did not ask students about high-stakes exams or any other assessments. My literature review reveals educational researchers know very little about what students think about the role assessments should play in their education. My study suggests that test results that are more detailed and delivered more quickly would please some students. I believe educators and administrators would also welcome detailed and timely information about student performance. I think the positive results from pushing for these changes would likely outweigh the costs and administrative hurdles such changes would require.

Quicker results and detailed feedback are changes my participants would like to see from high-stakes testing. I think Unger's findings suggest there are likely other types of information students would care to gain from the assessment process as well. Researchers could utilize surveys, interviews and focus groups to help reveal this information.

Motivation

All the parents I know have at least occasionally struggled with lack of motivation from their children. Many have also been surprised to find their children motivated by ideas or activities they would not have thought would spur their children towards action.

The same child who balks at completing homework and washing dishes may surprise his parents by taking a sudden interest in reading or in learning to cook.

It was a surprise to me that five of my participants expressed that they were generally motivated to do well on their CSAP exams. Reflecting back, I should not have been surprised given the importance parents and schools place upon these exams. It makes sense that pressure to achieve would have an influence on students' desire to do well on the tests. One of my participants, though, was clearly not motivated to do well on the exams; an attitude some participants felt was widespread among their peers.

My results lead me to believe motivation would increase if students were given a reason for taking high-stakes assessments. Something my participants said would have increased their motivation was quick and specific feedback regarding their test performance. They said they would like detailed information about how they could improve their future testing performance. The CSAP exam gave them too little information, too late, to be valuable.

I think shorter test sessions would decrease test fatigue and increase performance among students who struggle with assessments. I think motivation would increase in these students if they felt they had the support they needed before the tests and the accommodations they needed during the test. Measuring growth rather than measuring their performance would also likely increase their motivation. A student who struggles in school may always fall short against the proficiency threshold but may improve compared to his or her previous years test scores.

Classical test theory holds that an observed score consists of a true score and error. According to the theory, if a test instrument were administered to the same

participant multiple times, the average of the scores would yield something close to a true score. The true score represents the test taker's actual ability (Gliner and Morgan, 2000, p. 312). If motivation is affecting performance then it becomes part of the error that prevents test administrators from measuring a true score for an individual or group of students.

I have observed the effects of poor test motivation in my own classroom. I have had many students who did not put much effort into their CSAP exams. Some kids do not seem to care about their test outcomes. Others likely want the testing experience to end so they rush through, not caring about accuracy. For many of my students, high-stakes tests do not measure their hypothetical true score because of their lack of motivation. I think we need more information about what motivates some students to try hard on tests and others not. To get at this I suggest research investigations of the variables I mentioned in the policy implications section above. These include timely feedback, test duration and the effect of appropriate accommodations for special needs students.

Test Anxiety

Anxiety is frequently present in parent-child relationships. Sometimes the anxiety comes from organic causes, originating from a predisposition towards anxiousness either in the child or the parent. Sometimes anxiety results from conflicts between the parent and child. Anxiety may also result when children feel they failed to meet the expectations of their parents. Of course there are also many student-parent relationships where anxiety is not a factor.

This was the case among my participants. Some students reported little or no anxiety when faced with high-stakes tests. It is likely that a portion of all students, particularly those who are not confident in their ability or those who are highly concerned about the outcome of an assessment, will be anxious in any testing situation. In some cases test anxiety is big obstacle to performance. Two of my participants, Audrey and Maria said they felt their thinking was blocked when they confronted certain test questions. In this way anxiety, like motivation, prevents test givers from measuring a true score from some of those tested.

I believe if we are going to put children in a position where they feel anxious and fearful then we should make every possible effort to reduce their stress. How can we do this? First, we need to give appropriate instruments to our students. English Language learners need enough time to become facile users of written language before they are expected to take tests in English. Second, we need to be sure students have the accommodations they need when taking tests. These include extended time, scribing, help with reading when reading is not being directly assessed and using alternative instruments when appropriate. Third, we need to teach students strategies to cope with their stress.

Teachers and administrators should embrace this last suggestion for two reasons. They will see anxiety reduction as a means for improving test performance and they will wish to see the children in their care experience less distress during the testing process.

A pilot study conducted by Bishop (2006) found benefits from introducing a program designed to reduce the emotional stress of at-risk students. Her results indicated that engaged and empathetic teachers were able to reduce the testing stress of their

students using the program. She found that the program was effective when it was administered by sensitive teachers who were given sufficient time for a thorough presentation. The pilot program combined positive self-imagery and specific strategies for focusing student attention on the task at hand.

In a study conducted by Fields (2006) cognitive and study skills interventions designed to reduce test anxiety were successful in reducing self-reported anxiety levels. In this study, however, those reduced anxiety levels did not translate into increased scores on a standardized test instrument. Those looking to reduce anxiety solely for the purpose of making testing a more comfortable experience will be encouraged by these results, while those worried primarily about test scores will not.

I believe we need more information about how anxiety specifically impacts the results of high-stakes testing, more information about how it affects the students who take the assessments and more information about anxiety reduction strategies. Based upon the responses from my participants, I believe it is likely the scores of many students are negatively impacted by anxiety but I am more concerned that many students experience hours of emotional distress as a result of these annual testing batteries.

Human Connections and the Unknowable They

Most children, even those who complain about the restrictions their parents place upon them, seek to maintain connections with their parents. The Commission on Children at Risk (Kline, 2008) explored this need for connection with authoritative figures. Their research led them to believe children are “hardwired to connect” with adults. They suggest there is a crisis among many children because they lack “connectedness” with important adults. The first connection most children make is with their parents. During

their development they seek important connections with other adults as well. My participants seemed to be looking for this sort of connection with the people behind the testing process.

This need for connection among children is consistent with my experiences with students. My elementary children seek to have a connection with me as a person and I think they perform better when they are working for someone they know rather than for someone they do not. It should not have been a surprise for me to find that the high school sophomores in my study also wanted to find a connection with the people responsible for the tests they took each year.

I mentioned in an earlier discussion of this thematic element that one of my participants said she worked harder when she knew the person who made and graded a test. I think it would be a worthwhile policy experiment to introduce students to the people behind the testing process. Marketing firms and advertising agencies understand the importance of introducing customers to those responsible for producing their products. In a recent example of this strategy Budweiser released a television advertisement introducing an employee responsible for recycling aluminum cans. General Motors employed a similar strategy several years ago when they launched their Saturn division. They featured their production staff in a series of print and television ads. These strategies connect a human face to a product. Imagine a similar effort from a state department of education. Real people would explain the rationale for giving the tests in print and video infomercials. Vignettes about some test designers and test graders would accompany these ads. It just might work to improve student buy-in. Perhaps if the people

behind the tests were revealed, students would feel more compelled to work harder on the tests.

In the unlikely event a state adopted my suggestion to introduce students to the people behind the scenes of high-stakes testing, then I would hope somebody would step forward to evaluate the effectiveness of those changes. Barring that, another plan would be to investigate how test scores are related to student familiarity with test givers. It would be instructive to see if students would display better performance when they knew the people responsible for designing and grading their tests. If the research shows such a relationship, then states may be willing to implement my suggestions.

Relationship to Silence

I think relationship to silence is an example of a generational difference. The increasing availability of personalized and transportable music has radically changed the habits of American youth. A nearly steady stream of music accompanies many during their waking hours. This contrasts starkly with the silent environment in which high-stakes testing is done. Test givers share a position of many of the students' parents – Important, thoughtful work should be done in a silent environment. It appears that this attitude is not shared by all the youth taking the test.

Administrators might be interested in understanding this phenomenon, especially if it was established that high-stakes test scores increased when students had access to music during the testing process.

What would this look like? Michael suggested a system where students would choose from a list of musical genres much as airline passengers do during flight. I think it

is possible to go one step beyond and allow students to pick the actual songs they wished to hear while testing. The way this would work is students would make their selections from a large database of diverse selections. They would be able to choose as many tunes as they thought necessary to fill the testing time period or the songs could repeat if they chose fewer. These tunes would then form a personalized library to be played randomly from a central computer in the testing room. Students would likely complain about the sanitary choices but would likely prefer them to the alternative, which would be silence.

This emergence of this element was a surprise to me. I think others will find it surprising as well. The topic poses many possible research questions. The most important being: Is this student response to silence during testing widespread? Others are: Does silence during testing affect scores? Would test scores increase if students were given the kind of choice I suggested above during testing? I believe these questions are among the most compelling and important implications of this study.

Thematic Elements Summary

A better understanding of the relationship students have with high-stakes testing may lead to reforms to the reauthorization of ESEA legislation. Regardless of whether someone is a critic or a supporter of the accountability model, an understanding of the way students respond to testing could lead to better tests or better administration procedures. Students may still respond to the tests as if they were some unwelcome dictate from a parent but perhaps the intensity of their response would be tempered by incorporating some of these measures.

A Panel Discussion

Akiro Kurosawa's film, *Roshomon* (1950), explores four interpretations of a crime committed in ancient Japan. Initially, it appears a samurai has been murdered and his wife raped. Authorities apprehend a notorious outlaw and accuse him of the crimes. His version of the crime differs substantially from the rape victim's and from the only other witness, a passing woodcutter. Because discrepant accounts of the crime fail to provide any clarity for the court, they seek another witness. A medium is summoned and the samurai's story is channeled through her. The resulting contradictory accounts challenge the court and the viewer to determine which version of the story is true and to question the very nature of truth itself.

Named for the film it references, the Roshomon Effect embodies the Post-Modern notion of multiple realities and rejects the Positivist view of a single, objective reality. Wolcott (1994) suggests social science researchers can evoke the Roshomon Effect as a strategy for visiting multiple descriptive accounts from different participants.

Following his advice I have elected to employ the Roshomon Effect to examine the phenomenon of high-stakes testing through the eyes of four observers representing diverse stakeholder groups. In chapter four I analyzed transcript data using personal narratives and content analysis. Earlier in this chapter I discussed the implications of my findings and suggested possibilities for future research. In this section I wish to examine the high-stakes testing phenomenon from a broader, holistic perspective by way of giving voice to four observers.

The first voice is that of a high school student, the second is that of a state legislator, the third voice is my own, and the fourth is that of a departed educational

theorist. Kurosawa creatively introduced the ghost of the murdered Samurai in his film to provide an otherwise missing insight. I have borrowed from his approach and have summoned a ghost to testify as well. However, since I thought it unscholarly to channel a ghost through a medium, I have chosen a different strategy. I have employed an actor to stand in for my highly esteemed guest.

To write the student perspective I sourced transcriptions from my interviews as well as the themes I identified in chapter four. I have composed this student panelist's narrative based upon the opinions and attitudes held by my participants. It is, of course, not possible to embody the opinions of all of my participants within this single narrative. The response I have created attempts to voice the concerns I heard from several participants.

Although the State of Colorado is responsible for creating and administering the CSAP, it is the Federal government that has required a high-stakes assessment from all states under No Child Left Behind (Abernathy, 2007). Members of Congress crafted this piece of legislation and when the Elementary and Secondary Education Act next comes up for review, it is Congress that will write the new law. For this reason I have written a narrative from the view of a United States congresswoman from Colorado. Her opinions, like those of the student panelist are a synthesis of the opinions held by many members of Congress. I sourced candidate web sites as well as the literature to inform the content of this narrative.

My personal response is based upon twenty-five years of teaching, preparing students for standardized testing, and observing students while they are testing. Added to this background are extensive reading, writing and research that I have engaged in as I

have conducted this investigation. Writing this personal narrative was a way of plumbing and expressing my own attitudes and learning on this topic. The reader should keep in mind the opinions stated are my own.

The ghost referenced in the final response is that of John Dewey. He is one of the giants in the field of educational theory. It seems we in education have alternately embraced and rejected his teachings over the course of the last 100 years. It is hard to imagine a ghost to invite to this discourse who would bring more stature or baggage than John Dewey. It is always risky to attribute words to someone who is no longer able to speak. It is riskier still to attribute speech to someone of the significance and stature of John Dewey.

Structure of the Responses

In order to enforce a sense of structure to the accounts and to provide parallel structure, I have set up the responses as formal replies to a set of questions presented to all four respondents:

Has high-stakes testing improved education?

Does testing give an accurate view of school and student achievement?

Is the test relevant to students?

Is high-stakes testing fair to schools with a high number of low-income, or non-native language populations?

What do you think about assessing penalties towards schools that do not do well on high-stakes testing?

How would you change the current accountability model?

For this section of the document I have chosen to suspend conventional APA citation methods. I thought it awkward and unnatural to imbed citations into the panelist's

narratives, so I elected to include a list of citations following the panelist's responses.

There are no citations for the student panelist since I used participant's transcripts as the source for her response.

The Discussion

A tall man with graying hair and carefully trimmed goatee strides confidently to a microphone placed in the center front of a stage in a large auditorium. Purple curtains provide a backdrop. Directly behind the microphone is a table containing a water pitcher and five tumblers. On the far side of the table five upholstered swivel chairs are arranged in a semicircle. The chairs are empty and dimly lit. The speaker acknowledges applause from the nearly filled auditorium with two half bows and then begins to speak:

Good evening, this is Wayne Sonorous. Welcome to this month's NPR Educational Forum. Our topic this evening is high-stakes testing. For reasons that will soon be apparent, it is appropriate that tonight's forum is being held in the historic John Dewey High School in Denver, Colorado. This beautiful building, built in 1912, was one of the schools in the early twentieth century that embarked upon a radical new pedagogical course inspired by the famous and influential educational theorist and reformer John Dewey.

A visitor to this school in 1912 would have observed students engaged in educational projects shaped by their own interests. This school was a laboratory for student-centered learning. Through the years student-centered learning has waxed and waned in this school and in American schools at large. A visitor today would still find students engaged in projects of their own creation but if she came here during the testing window in March she would see sophomores taking the Colorado State Assessment Program exams. For a total of 12 hours, over the course of two weeks, sophomores in

high schools all over the state take state-mandated exams upon which the fate of their school rests.

It is these high-stakes tests that are the topic of tonight's Educational Forum. Four speakers will respond to questions chosen from among the many we received from our listeners. Each speaker was given the questions in advance. He or she will be given 12 minutes to respond to the questions.

Three people emerge from behind the curtain on the left side of the stage and stand in front of the chairs behind the speaker. The first to enter is a short girl with freckles, green eyes and fiery red hair. She teeters a little on her heels and tugs at her dress as she stands on the stage. The second is a middle-aged woman slightly taller than the girl. She has blonde, carefully coiffed, and totally immobile hair. She is dressed in a blue pantsuit with a red-white and blue scarf tied around her throat. The third person is man, taller than the two women but shorter than the announcer. He is middle-aged, bald, with a close-cropped beard. He is dressed in a grey suit and a pair of worn Birkenstocks.

I would now like to introduce tonight's speakers. Our first speaker, Farrah Rubin, is a student from Silverheels High School in Ute Creek, Colorado. She is a sophomore who recently completed her last state mandated tests (*Farrah waves to the crowd as the audience applauds enthusiastically*). Our second speaker, Connie Campaynor is a Republican United States congresswoman from Northern Colorado who was recently reelected for her third term last fall. (*Polite applause*) Our third guest this evening, Tim Pearson, is a veteran teacher and graduate student who like Sarah, is from Ute Creek, Colorado. (*Moderate applause*)

Our fourth guest will join us a little bit later. We will keep his identity secret until he appears. I assure you he is a man of tremendous stature in the educational community and one who is quite used to the dissenting opinions his views often generate.

The three guests settle into their seats, Wayne Sonorous remains standing and begins to address the audience.

Our first speaker is Farrah Rubino. Farrah, you will have twelve minutes to speak.

Farrah stands and takes Wayne's position at the microphone in the center of the stage. She pushes her thick hair aside, smiles wryly, and begins to speak.

Hi! I am a little bit afraid to be speaking to you today but I am also glad because people don't often ask kids what they think of the CSAP I mean high-stakes testing. In Colorado our state tests are called CSAPs but everybody has been reminding me to call them high-stakes tests tonight.

I have been taking the CSAPs since fourth grade. This year I took my last one. I am really happy about that. I didn't really think about the tests too much until I started Junior High. My friends and I started talking about how stupid they were. I thought they were pointless also but I didn't really think too much more about them until Mr. Pearson interviewed me for his dissertation. Yeah. Then I started thinking more about why they were stupid and pointless. So that is why I am here. NPR gave me some questions to answer and I have some notes here but sometimes I get all hyped up and forget to look at them (*laughs from the audience*).

Has high-stakes testing improved education? I don't think so because I think the testing has messed up education. This year we spent like three days testing. That means that for three days we didn't go to our classes for half of the day. And that is just the testing itself. There is all the time used in preparation for testing. The teachers are always saying something like, " You know this might be on the CSAP exam." Or, "We better review some algebra before you have to take your CSAPs next week." I mean we get tired of hearing that stuff. Yeah. It's kind of weird.

So, my answer would be no, high-stakes testing hasn't improved education because the teachers are spending time getting us ready for the test when they could be teaching us new things and we waste time we could be learning by spending so much time testing. I think we would learn more if we didn't have the tests at all.

Does testing give an accurate view of school and student achievement? I didn't pay too much attention to the testing when I was in elementary school but as a junior high and high school student I started to think more about testing and to notice what other students were saying about it. I started to think about the test, you know, in eighth and ninth grade. I became more critical about it. I still tried hard, though, because I always try hard but I don't think a lot of other kids try hard. I have taken CSAP tests with kids who just mark any answer they feel like and then go to sleep. I mean, I don't think it is right that they do that and I wouldn't do that but I understand, you know, why they do it. They are just sort of tired of all the testing and they know it doesn't matter.

I know lots of kids that feel this way about the CSAP exams. Some say the test isn't important, some just hate being tested, and some get test anxiety. I know students who feel like they are going to throw up when they take a test and they can't think very well. They can't remember stuff they really know. And then there are students who just came from Mexico and they are still learning to read and write in English. Some schools have lots of those kids and I don't think it is fair to compare those schools to schools with lots of kids that speak better English.

So no, I don't think the tests do a good job of telling about school or student achievement. Too many kids aren't motivated or don't do well on the tests. I just don't think the tests tell that much about schools or kids. Maybe the tests in elementary schools

are better because I think the kids take it more seriously but a lot of kids I know just don't try that hard.

The next question is: Are the tests relevant to students? I am so glad I get to answer this question. I so don't think the tests are relevant to students. First of all we don't get the results back from the CSAPs until the next school year and when we do get the results they don't tell us how we could improve for the next test. They just tell us if we are proficient or advanced. We don't get any specific information about what we missed.

Next year I will take the SAT and ACT exams. Those are important to me because if I do well the scores will help me get into college. I will try my hardest because I want to get into a good school. But the CSAP, it doesn't matter to me. I mean, I want my school to look good and I want my teachers to look good but it doesn't help me personally. My grades stay the same if I don't do well and I've heard that colleges don't care about the scores. I have always done well on the CSAPs but it hasn't helped me at all.

I know that most of my friends think the same way about the test not being relevant. When we were younger we wanted to do well on the test. I don't know why, maybe just because we wanted to please our teachers. Now everybody asks, "why should I try so hard?" We have all figured out that it doesn't help us or hurt us.

The next question is about whether high-stakes testing is fair to schools with lots of minority or low-income students. I don't think so because those schools have lots of kids from Mexico and they are still learning to read and write in English so how could they do well on the CSAPS. Yeah. And, you know some of those kids are from homes

where they don't have parents that bother them about their homework like my parents always have. That's just what I think. I can't think of anything else to say about that question.

What do you think about assessing penalties towards schools that do not do well on high-stakes testing? I think they should quit punishing schools that have low scores. I learned in my psychology class about how punishment doesn't work very well, like when parents are trying to make their kids behave. But that seems just what they are doing to schools. If they don't do well, they get punished. I know there are schools here in Denver where they took over the school or closed them down and the kids had to go to a school they didn't know just because they didn't have good test scores. I think they should help those schools get better, not close them down. I wouldn't like it if they made me go to a different school.

The next question is a hard one for me to answer. How would you change the current accountability model? My grandma always says, "Don't criticize something unless you have a better way to do it." Well, this question made me do like my grandma says. I had to think up a better way. *(Soft laughter from the audience)*

I think the first thing they should do is make the tests shorter. I spent 13 hours on the CSAP tests this year. That's just crazy! I think they should be, like, two or three hours max. I also think the teachers shouldn't be allowed to teach to the test. I mean, I think they should teach us the stuff they think we should know, but they shouldn't teach us stuff just because they know it will be on the test.

Another thing (and this is one I think adults just don't understand), it is really hard for us to sit in a quiet room for a couple of hours. I mean it isn't natural. Almost

everywhere I go there is music, or at least sound, but in the room when we are testing it is weird, it's like dead, like there is no energy left. I think they should figure out some way so we can listen to music if we want, you know, like when you ride on an airplane. They could have different channels, like classic rock, rap, country, alternative, or metal. That way we could listen if we wanted, but if there were kids who didn't want to, they could do their tests in silence.

So, that's what I think about high-stakes testing. Thanks for letting me talk about it. I know I am just one kid but I know lots of kids who think the same way I do about these tests. Well, thanks and have a nice night.

The audience applauds warmly as Sarah takes her seat and Wayne takes the microphone.

Our next speaker is United States Congresswoman Connie Campaynor who represents Northern Colorado in Washington. Please welcome her to tonight's forum. *Polite applause accompanies Mrs. Compaynor's movement from her chair to the podium.*

Thank you. Thank you as well to National Public Radio for inviting me to this forum.

Before I begin responding to the questions provided by the panel I would like to commend Farrah for her thoughtful and well-delivered response. Great job, Farrah!

My first response is to the question: Has high-stakes testing improved education? My answer is yes, and no. Sounds just like a politician you might say. Please let me explain. It is difficult to separate high-stakes testing from the whole of No Child Left Behind legislation. I would first like to address the strengths of the law as it relates to assessment. The legislation has reporting provisions that have forced schools to give

much more information to parents about how well schools are educating kids. This has done two things. It has forced schools to improve teaching by focusing on what is really important and it has given parents much more information about what happens in their children's schools.

This leads us to another positive result of No Child Left Behind and testing. There are provisions that help students in poor-performing schools. If schools do not perform well on the tests, there are mechanisms in place that require them to provide tutoring programs for their students. If the school's test scores still do not improve, the law allows parents to move their children to other schools of their choice. These changes have given parents much more influence over their children's education.

Another positive result of school accountability is that education has improved for many students. The Fordham Foundation investigated whether states were making educational progress. They concluded 20 states had seen dramatic improvement in the number of students whom were "proficient" or better according to the test results. To do this they compared state test scores to student scores on the National Assessment of Educational Progress results. This is evidence that some children in some states are getting a better education. The bad news is that for children in 30 states there is no evidence of improvement. To see improvement for all students we will have to have high expectations and high thresholds for improvement in all states.

A problem with the current system is that although parents may elect to move their children from failing schools, their choices are limited. If the test scores are low in a particular school parents can chose to send their children to other schools but because of the way states have implemented No Child Left Behind, those schools are almost always

other public schools. If parents truly had access to their tax dollars they would be able to send their children to private schools as well. Then we would have real choice.

Does testing give an accurate view of school and student achievement? As I mentioned in my response to the first question, for the students in 20 states, the answer is yes. The students and parents in those states are fortunate enough to have rigorous standards and tests that drive effective educational practices. Students in other states may be lucky enough to attend a school that exceeds those standards and has excellent test scores. They may, on the other hand, be unlucky and attend a school in a state that has set the bar too low. In that case the test scores may not accurately report the lack of academic rigor within the school or school district.

Are tests relevant to students? Farrah made it clear high-stakes tests were not important to her. I have talked to other students who said they appreciate the tests because they show their schools are doing a good job of providing them with a quality education. The solution to making the test relevant to all students is to tie testing to high school graduation. Some school systems are requiring seniors to take an exit exam before they receive their diploma. Those who fail either need to continue with high school or accept a different type of diploma. I think if diplomas were linked to testing, students would understand the importance of doing well on the tests.

Is high-stakes testing fair to schools with a high number of low-income, or non-native language populations? No Child Left Behind was crafted with low-income, minority, and non-English speaking children in mind. For too long we ignored the needs of these populations. One of the things this legislation does is require reporting of scores from many sub-groups. In the past it was possible for a school with a high number of

high-achieving students to ignore their minority, low-income, or special needs students and still get good average test grades. Now a score from each subgroup must be reported separately. In this way we should truly leave no student behind.

That said, I think there is room for improvement in the testing process. It is not perfect. We need to do a better job of making sure students with disabilities and English language learners have a fair opportunity to demonstrate their knowledge when they are testing. Sometimes these groups of students are not getting the testing accommodations they need.

What do I think about assessing penalties towards schools that do not do well on high-stakes testing? I think the penalties have been instrumental in bringing about educational reform. The threat of losing funding or the threat of a school takeover has forced many low performing schools to turn around. It is remarkable what a little accountability has done for many schools.

Of course this has not been universal. Some states have made the tests so easy they do not truly show how poorly some students within those systems are really performing. In these states many schools have not had to face the music. They have not had to respond to the threat of the penalties. We need to work with states to make sure each of them have rigorous standards for achieving Adequate Yearly Progress.

What changes would I suggest be made to the current accountability model? I have mentioned a few things already. I will summarize my suggestions. First, I think we should give parents real choice for educating their children. We need to make sure charter schools can be formed when parents feel there is a need for an alternative educational model. We also need to give parents other opportunities to educate their children in ways

they see fit. This means allowing them to take funds already earmarked for their children's education and use them to enroll their children in any school they wish, public or private. And to further provide families with choice, we need to devise a way to use those funds to help parents provide home schooling for their children. Second, we need to work with states to make sure all of them have rigorous standards and tests that adequately assess the teaching of those standards. Finally, I think we need to be sure testing is fair for students with mental or physical disabilities and for students who are English learners. I think these changes will make No Child Left Behind come closer to fulfilling its intent. Thanks so much for giving me the opportunity to speak this evening.

Polite applause ensues as Mrs. Campaynor takes her seat and Wayne Sonorous approaches the microphone.

Our next speaker is teacher and graduate student, Tim Pearson

(Warm applause)

The bald man in the grey suit approaches the microphone.

I would like to thank NPR for inviting me to this forum and the audience for joining us for this event.

I have recently completed a research study investigating how students respond to high-stakes testing. This research has shaped my thinking about testing. I welcome the opportunity to speak tonight because it has forced me to sort through my opinions about educational accountability and synthesize them for this panel.

Has high-stakes testing improved education? I suppose it depends upon what one means by improved. A recent Rand study, an extensive survey of teachers in three states, provides some perspective on this question. If by "improved" you mean have public school students become better test-takers? Then the answer is yes. If by improvement you

mean are teachers more focused on educational standards? Then the answer is yes again. If by improvement you mean are teachers searching for and using innovative teaching methods as a result of testing pressure? Then the answer is also yes.

The same study, however, highlights some possible problems with the new accountability structure. One of the findings was that teachers reported spending more time on test preparation. Another was that teachers felt educational content had been narrowed. Teachers said they felt pressure to focus on students who were almost proficient on tests at the expense of those students who were so far below proficiency they were unlikely to catch up. Teachers noticed a lack of consistency between their district standards and material tested. Teachers also reported their morale was lower as a result of testing pressures.

The results of other studies reported mixed results as well. For example, a large-scale study conducted by the Center on Educational Policy found many states had increases in math and reading scores at the elementary and middle school level. The same study reported only a few states had increases in math and reading at the high school level. Another study analyzing the scores of over 300 thousand students found initial improvements in reading and math scores but later testing showed a decline in scores.

The results of these studies are ambivalent at best and the Rand study provides some disturbing information about how high-stakes testing is shaping instruction. I believe these studies support my opinion that this kind of testing has done more harm than good.

Does testing give an accurate view of school and student achievement? One of the most obvious problems with the current testing model is that it relies on a single annual

test. The Standards for Educational and Psychological Testing address this issue explicitly. They state high-stakes outcomes should not be linked to a single test. One test gives a mere snapshot of student performance. It does not provide an accurate measure of the learning test designers intend to assess.

I believe we fail to test student behaviors that will create reflective future citizens. We do not test for compassion, honesty, collaborative skills, curiosity, or problem-solving strategies. These are among the most important skills students will need to be contributing members of the 21st century. Instead, we continue to assess skills that are discreet and easily assessed. Many of these skills are important but they represent only a subset of a larger set of knowledge and skills that students need to become valuable citizens.

Another problem with high-stakes tests is they do an excellent job of measuring a school characteristic they were not intended to measure - the affluence of a school's population. Researchers have found a strong relationship between affluence and test performance in large-scale studies conducted in Minnesota and Ohio. I think this pattern would be found in nearly every state and school district in the country. If this question was rephrased to ask, "Does testing give an accurate view of the number of students eligible for free or reduced lunch in a particular school?" Then the answer would be, "Yes."

Is high-stakes testing relevant to students? My research indicates students have difficulty connecting high-stakes testing to their own future. Farrah voiced this effectively in her response to this question. I would like to add a single point to Farrah's response. Congresswoman Compaynor suggested making the tests relevant to students by

linking them to high school graduation. This may prove to be motivating for some students, but I think it could lead to many undesirable consequences as well. If this practice was widespread I think we would see more instances of test anxiety, increased dropout rates, and more frequent examples of cheating or gaming. I think it is also likely teachers would spend more time teaching to the test and teaching test-taking strategies. Linking graduation to high-stakes tests might succeed in amplifying many of the negative consequences of high-stakes testing we already see.

Is high-stakes testing fair to schools with a high number of low-income, or non-native language populations? No, there are a number of reasons No Child Left Behind deals an unfair hand to schools with these populations. The first, and most glaring in my mind, is the Adequate Yearly Progress targets set by the legislation. If a group of students begins with knowledge and skills well below the mean, their schools must generate test scores for that group that show more than a year's growth in order to meet proficiency targets. The National Education Association has joined others in calling for high-stakes testing to measure yearly growth rather than using the same yardstick for all students. In this way schools with a high number of low-income or non-native language students have an equal chance of showing growth as schools that have more affluent and less diverse populations.

A second problem lies with the designation of subgroups by No Child Left Behind. The intention is to force schools with high test scores to pay attention to minority and special needs students whose scores might otherwise be offset by those from the rest of the school's population. Unfortunately, the system for reporting subgroup performance does not work as it was intended. A researcher examining data from

Minnesota schools found the higher the number of subgroups a school reported the more likely they were to fail Adequate Yearly Progress goals. Larger schools are more likely to have multiple subgroups and therefore have more opportunities to fail AYP. Thus the current system favors schools with lower student populations and less diversity.

A third problem with No Child Left Behind is with English Language Learner (ELL) requirements. Researcher Ronald Solorzano describes some of the difficulties that arise when high-stakes assessments are used with bilingual students. One problem he identifies is incomplete language fluency. Students are frequently tested in English before they are ready. He also found both English and Spanish versions of high-stakes tests contained culturally biased test items and that the Spanish tests were not correctly normed. Solorzano's findings suggest the methods for testing of ELL students warrant critical examination.

High stakes testing as it is currently structured establishes unrealistic learning thresholds, penalizes schools for having diverse populations and fosters invalid testing of English language learners. Because of these problems it is not fair to schools with large numbers of low-income and emergent English speaking students.

What do I think about penalizing schools that do not do well on high-stakes testing? A widely accepted premise of modern education is that punishment is ineffective when used with students. Farrah mentioned this in her response. The more authoritarian the response, the more ineffective the result when applied towards children in classrooms. Problem solving strategies where teachers work with students to find solutions to misconduct have proven to be more effective than punitive measures. This is not to say punishment is absent from our schools or that people have stopped believing

punishment is an effective motivator for behavioral change. The widespread use of punishment and belief that punishment is effective are likely reasons legislators who crafted No Child Left Behind thought penalties would be an effective means for motivating school districts to improve.

Punishment does not work well with children and I do not think it works well with institutions. Strict penalties lead children to find ways to avoid consequences without making desired changes. A student may lie to avoid punishment or find other means to avoid it. The same thing has happened in schools. Two Texas researchers reported increased retention rates for ninth grade students in that state after high-stakes testing had begun. This is important because holding back ninth graders prevents low-performing students from taking the required Texas assessment as tenth graders. Many of these retained students later become dropouts. There are other examples of teachers and administrators intervening in dubious ways to influence student performance on high-stakes tests as well. I think cheating and gaming strategies are an inevitable consequence of assessing penalties on poor performing schools.

Another important point I wish to make about testing penalties is that they are directed disproportionately towards schools that serve poor and minority students. Teachers and administrators in affluent and low-minority schools get a pass on penalties, while staff in high-minority, low income schools face funding loss, closure and restructuring. Among other problems, this creates pressure for teachers, and likely administrators, to look for new assignments in schools where the student population scores well on tests. This results in the siphoning of talented and experienced teachers from the schools that most need their talents.

I think we need to offer support rather than assigning penalties to schools having difficulty meeting AYP. Support should include programs designed to attract and retain talented teachers and grants directed towards increasing monetary and physical resources in poorly funded schools.

How would I change the current accountability model? First of all I want to say I do not think the Federal government should be involved in accountability testing. I believe states and school districts should have more autonomy in these matters. That said I recognize that once the Federal government becomes involved in a process, it is unlikely to retreat from it. Accepting this possibility, I will suggest a few changes.

The most substantial change I recommend is to measure student growth rather than assessing student performance based upon an absolute proficiency threshold. I discussed the reasoning behind this earlier in my discussion. Another change I have suggested is to remove penalties and replace them with programs that support struggling schools. It is time to stop beating up on “failing schools” and give them assistance to overcome their complex problems.

A big change I would like to see if we continue to use mandated assessments is a shift to using multiple, instead of single, assessment instruments. Two or three assessments of shorter duration would likely provide more valid data than a single lengthy test. A possible side benefit of this change would be reduced student fatigue and increased motivation.

The next change I suggest goes back to something I said in response to the previous question about the importance of the information high-stakes tests provide. I think the assessments should be designed to display student problem-solving strategies.

This would encourage the teaching of these skills in the classroom. It would be necessary to do this in a way that was culturally sensitive, and in a way that was open to multiple problem-solving strategies. By this I mean that test graders would be asked accept multiple solutions that displayed critical thought. A shift towards assessing critical thought would be an improvement but tests would still not measure other valuable attributes such as compassion and collaboration skills. It is, though, a step in the right direction.

Finally, I think the government should fulfill its promise on two points. The first is the promise imbedded within the introduction to No Child Left Behind to provide flexibility to states and school districts. Federal officials may argue they have been flexible because they have allowed different states to use different testing instruments. That is true, but there has been less flexibility about testing timetables, penalties directed at failing schools, and subgroup designations. The second failed promise was to provide the funds needed to implement federal mandates. Just a few months after the legislation had been signed, two key sponsors of the legislation, Edward Kennedy and George Miller, publicly issued a letter accusing the Executive branch of renegeing on its promise to fund No Child Left Behind related expenditures. Following their call for more spending, nine school districts joined the National Educational Association in suing the U.S. Department of Education on the grounds insufficient funding was provided to implement the legislation.

No Child Left Behind is a complex piece of legislation. This evening we have only addressed the accountability component and we have not given even that portion the full attention it warrants. I encourage the audience to look critically at how the legislation

affects their children and their schools and then make up their own minds about whether the legislation has been successful or not.

Thank you and good evening.

Warm applause follows as Mr. Pearson takes his seat and Mr. Sonorous steps up to the microphone.

I promised you we would have a mystery guest of great stature speak this evening. Dr. William Davis, renowned educational scholar, began making public appearances in the 1980s in the guise of the great school reformer, John Dewey. Two years ago he retired from his teaching position at Sinclair College in Chicago. He now tours the country making appearances at schools, colleges and town halls as John Dewey.

Please welcome Dr. Davis appearing as John Dewey!

Loud applause ensues as Wayne Sonorous greets a tall thin man with thick gray hair dressed in an antiquated grey flannel suit. Dr. Davis takes the microphone as the applause continues.

Thank you. Thank you so much for the warm welcome.

I appreciate this opportunity to bridge several decades and join you as a recognized critic of educational practices past. Many of us have the pleasure of critiquing the policies of our own time; few of us are accorded the opportunity to criticize the policies of the future. I gladly accept this challenge to speak to stakeholders and policymakers of this new century.

The first question I will address this evening is: Has high-stakes testing improved education? Let us assume, merely for the purpose of our discussion tonight, that every student in this nation was able to perform satisfactorily on the assessment instruments employed by various states. This performance would demonstrate to the country that

these students had mastered the facts and skills deemed necessary to be properly educated. This state of affairs would likely make many lawmakers and educators content.

But would their performance mean these students were well prepared for the future? Would these children be able to think in a reflective manner? Would they be able to effectively analyze an ethical situation and arrive at well-reasoned conclusion? Would they be able to think in a scientific way? Would they have an appreciation of democratic values? Would they have an aesthetic sense? In short, would these students be ready to become productive and participatory members of society, or would a narrowly defined and partially conceived form of education merely prepare them to be competent takers of tests?

I do not believe high-stakes testing has improved education in this country or that it is even capable of improving education. But I will address that more fully in response to the next question. Which is: Does testing give an accurate view of school and student achievement?

Like archers taking aim at their targets, teachers and educational leaders have fixed high-stakes tests in their sights. Their focus is misdirected. I think many of them believe that achieving excellent test scores is the most important role of education. When educators direct their efforts in this way, they often do so at the expense of the rich experiences that should be cultivated in classrooms. With tests as their focus, teachers tailor lessons to improve students' test performance. Teachers attempt to pour knowledge into the minds of their students as if they were empty vessels that needed filling. Effective education should not be concerned with offering up more content; it should be directed at providing more educative experiences. These experiences should be carefully

tailored to encourage students to become reflective citizens. We cannot know what discreet skills and bits of knowledge the citizens of the future will need; therefore, the best course is to provide them with thinking strategies that will allow them to make intelligent decisions when they encounter the problems of the future. It is hard for me to conceive of a standardized test that would allow educators to assess whether students will become tomorrow's successful citizens.

Are the tests relevant to students? I can add little to what the young lady who spoke previously had to say about this subject. She described how she was unable to establish a connection between high-stakes tests and her educative needs.

Many of my critics have falsely claimed that I place the child above curriculum as the most important element in education. In truth, I have frequently said I value both the role of the individual student and educational content in the educational process. In this instance, regarding the relevance of the high-stakes testing, I favor the position of the students because I do not believe the tests serve to further important curricular content. Instead, testing appears to be designed to favor the desires of political leaders rather than the needs of students. It is difficult, given this role of high-stakes tests to see how they can be relevant to the lives of students.

Is high-stakes testing fair to schools with a high number of low-income, or non-native language populations? I have frequently spoken against discrimination directed towards minorities and immigrants in schools. I believe it is in the national interest to create a culture in our schools that encourages the acceptance of people from a wider community. For this reason I was pleased to hear of the advances that have been made in

educating minority and immigrant students in many of your public schools during the last three decades of the previous century.

Unfortunately, some of this progress is being undone. I think accountability testing has played a role in this process by emphasizing test scores. Many parents believe test scores reflect the quality of education delivered in a particular school. Concerned by lower scores in schools with large numbers of minority and non-English speaking students, parents frequently choose to send their children to schools with better test scores but less diversity. In this way, advances made in the late 20th century towards school integration have been diluted by the flight of affluent White students from previously integrated schools. A strong democracy requires children in public schools have equal opportunities to learn. This is unlikely to occur when race and class separate children.

Segregation is one problem linked to high-stakes testing; a lack of cultural sensitivity is another. The current accountability system encourages teaching of discrete skills and facts. When a child from a low-income background enters school, educators and administrators often see a child lacking in many of these traits. They view the child from a perspective of deficiency, seeing what is missing, rather than seeing the rich cultural understandings and experiences the child brings with him. The culturally different child comes to school confident in the learning he has acquired at home and in his neighborhood but soon comes to understand he is lacking in the technical learning that is expected of him at school. He has trouble bridging the gap between these two worlds. Soon he begins to feel isolated at school. For this reason I think the real poverty found in many of your schools is a poverty of understanding, a poverty caused by a lack of sensitivity to the needs of students.

Because high-stakes testing exacerbates school segregation and because it encourages schools to look at culturally different students from a deficit perspective, I think it has not been beneficial for low-income and minority students.

The next question is: What do you think about penalties applied towards schools that do not do well on high-stakes testing? To be perfectly blunt, I think they are a poor idea. I am a little surprised, and frankly disappointed, to see educators in this new century still applying punitive measures. It was commonplace in the latter part of the 19th century and well into the last century to punish students for their lack of interest and inattention. On numerous occasions I spoke against this practice. I explained to educators that student attention would come when they established a meaningful connection with the content, not from creating a climate of fear in the classroom. Just as I think punishment of students for their inattention is not appropriate or productive, I believe it is inappropriate to punish schools for poor performance.

I particularly think it is inappropriate to penalize schools based upon the narrow content high-stakes tests measure. Rather than revealing if students are prepared to be reflective thinkers capable of becoming valuable contributing members of a democratic society, these tests describe how well a particular group of students has been trained to take tests. Penalizing schools for failing to arrive at these misdirected goals is as inappropriate as punishing a student for inattentiveness.

What changes would I suggest be made to the current accountability model? Mr. Sonorous indicated at the outset of this discussion that each panelist had 12 minutes to respond to the questions provided. I did not check my watch when I began, but I am

certain I have far less time left than I could spend on this question; therefore, I will attempt to keep my responses uncharacteristically brief.

The time when I was conducting studies in the Lab School at the University of Chicago was a period of great change in the world. New modes of electronic communication fostered easy and rapid interaction. You are currently undergoing similar substantial changes in electronic communication. Fast and accessible electronic connections have given you access to information in ways we could not have anticipated 100 years ago.

I urged the educational leaders of my day to create school environments that would produce citizens who would be ready to accept the challenges posed by a rapidly changing environment. I urge your current educational leaders to do the same. The schools of my era were structured upon a medieval conception. In many ways your schools resemble those I wished to see restructured. Then, as now, schools were largely concerned with the accumulation of information and with gaining control of symbols. What I desired schools to do then, and what you should desire from them today, was to produce future citizens who possessed the desire to create and produce for purposes of both utility and art. If you conceive an educational system that values student interest over the control of symbols and facts, you will produce members of society who are more vital and more sustaining of culture.

The rapid changes I mentioned are unsettling for many. One frequent consequence resulting from the discomfort caused by change is that people resort to dogmatic thinking. Leaders propose, and citizens are willing to accept, prescriptive and narrow rules and regulations that characterize limited thinking during times of significant

change. The educational accountability model we are discussing today is an example of just such dogmatic thinking. The future requires flexible thinking from reflective citizens, not entrenched, and redundant strategies.

I suggest you abandon such dogmatic thinking and create a system that measures learning by observing the types of thinking students in your schools are capable of demonstrating. Can students conduct an investigation using scientific inquiry strategies? Do they ask thoughtful questions? If you must test children, test whether they can do these things, not whether they can provide a prescribed answer to an either-or question. You need a nation of life-long learners capable of contemplating the issues of the future, students who are able to embrace rapid changes in communication and technology, not dogmatic thinkers.

I thank you for this rare and unusual opportunity to visit an era that is much changed, yet in many ways, quite similar to my own time. Before I depart, I wish to encourage those among you whom are advocating for educational change to continue your efforts but I warn you against accepting my view or anyone else's view of educational reform without looking at it with a critical eye. My educational ideas may guide you, but you should also hold them up to what you observe and what you learn, and then make your own way forward.

Good night and thank you for this opportunity.

Enthusiastic applause follows as Wayne Sonorous shakes Dr. Davis's hand and strides to the microphone.

I would like to thank tonight's panelists for their reflections and our audience here at John Dewey High School in Denver, Colorado. Join us next month in Saint Paul,

Minnesota when we look at college entrance exams on Education Forum. This is Wayne Sonorous for NPR saying good night.

Citations for Preceding Panel discussions

Connie Campaynor

Brown (2007). Congresswoman Marilyn Musgrave (n.d.). John McCain on Education (n.d.). Klein and Hoff (2008). Lips (2004). No Child Left Behind (n.d.). Washington (n.d.). What is compassionate conservatism? (n.d.).

Tim Pearson

Berliner (2006). Becker (2003). Cronin, Kingsbury, McCall and Bowe (2005). Grow your own teachers: An Illinois initiative (n.d.). Poetter, Wegert and Haerr (2006). Hamilton et.al. (2007). Gulek (2002). Kennedy and Miller (2002). Kober, Chudowsky and Chudowsky (2008). Martinez and Martinez (2002). NEA: NEA's Positive Agenda for the ESEA Reauthorization (n.d.). Sacchetti (2005). NEA: court hears NEA appeal in NCLB unfunded mandates (2006). Solorzano (2008). Tulley and Chiu (1995).

William Davis (as John Dewey)

Benson, Harkavy and Puckett (2007). Brickman (1975). Dewey (1980). Dewey (1916). Handlin (1959). Simpson (2006), "*Some of John Dewey's writings about*" (n.d.).

REFERENCES

- Abernathy, S. (2007). *No Child Left Behind and the public schools*. Ann Arbor: University of Michigan Press.
- Baylis-Heerschoop, C. Federal aid to education. *Notre Dame*. Retrieved February 7, 2008, from <http://www.nd.edu/~rbarger/www7/fedaid.html>
- Becker, J. (2003). Texas miracle - zero dropout rate. Retrieved November 13, 2003, from <http://www.mathforum.org/epigone/mathed-news/trimpfrathol>
- Beilock, S. L., & Carr, T. H. (2004). When high-powered people fail: Working memory and choking under pressure in math. *Psychological Science*, *16*(2), 101-106.
- Benson, L., Harkavy, I., & Puckett, J. (2007). *Dewey's Dream: Universities and democracies in an age of education reform*. Philadelphia: Temple University Press.
- Berliner, D. (2006). Our impoverished view of educational research. *Teachers College Record*, *108*(6), 948-995.
- Bishop, N.S. (2006). Implementation and assessment of a test anxiety reduction program presented to 10th graders and their subsequent performance on the MontCAS Criterion Referenced Test. *ProQuest* (No. AAT 3244369).
- Blair, J. (2001). Testing opponents speak out in ESEA homestretch. *Education Week*, *20*(43).
- Brickman, W. (1975). *John Dewey, master educator*. New York: Greenwood Press.
- Brown, J. (2007, Mar. 5). *News from the Republican Party and elected officials*. Retrieved June 12, 2008, from <http://www.larimergop.org/inthenews.htm>
- Carter, S. C. (2001). *No excuses*. Washington DC : The Heritage Foundation.
- Clandinin, J., D., & Connelly, M., F. (2000). *Narrative inquiry*. San Francisco: Jossey-Bass.
- Clarke, A. (2005). *Situational Analysis: Grounded Theory After the Postmodern Turn*. Thousand Oaks: Sage Publications, Inc.

- Congresswoman Marilyn Musgrave*. (n.d.). Retrieved June 12, 2008, from <http://musgrave.house.gov/Issues/Issue/?IssueID=1641>.
- Connelly, M., F., & Clandinin, J., D. (1990). Stories of experience and narrative inquiry. *Educational Researcher*, 19(5), 2-14.
- Coppola, E. M. (2005). The American dream and the purpose of education in a high-stakes policy environment. Retrieved August 5, 2005, from www.aera.net/uploadedFiles/Meetings_and_Events/2005_Annual_Meeting/Program/174-201%20Wednesday%20Morning.pdf
- Colorado Department of Education (n.d.), Unit of student assessment. Retrieved September 24, 2008 from http://www.cde.state.co.us/cdeassess/documents/csap/csap_plds.html
- Creswell, J. W. (1998). *Qualitative inquiry and research design*. Thousand Oaks, CA: Sage.
- Davison, M., L., Young, S. S., Davenport, E. C. J., Butterbaugh, D., & Davison, L., J. (2004). When do children fall behind? What can be done? *Phi Delta Kappan*, 85(10), 752-761.
- Dewey, J. (1916). *Democracy and education; an introduction to the philosophy of education*. New York: MacMillan.
- Dewey, J. (1963). *Experience and education*. New York: Collier Books.
- Dewey, J. (1980). *The school and society*. Carbondale, Illinois: Southern Illinois University Press.
- Fields, E. (2006). Test anxiety interventions for middle school students. *ProQuest* (No. AAT 3245535).
- Fine, L., 20(38) 24-26. (2001). ESEA, minus vouchers easily passes House. *Education Week*, 20(38), 24-26.
- Fowler, F. C. (2004). *Policy studies for educational leaders*. Upper Saddle River, NJ: Pearson.
- Gergen, M. & Davis, S.N. (2003). Dialogic pedagogy: Developing narrative research perspectives through conversation. In R. Josselson, A. Lieblich & D.P. McAdams (Eds.), *Up close and personal: The teaching and learning of narrative research* (239-258). Washington D.C: American Psychological Association.

- Good, C., Aronson, J., & Inzlicht, M. (2003). Improving adolescents standardized test performance: An intervention to reduce the effects of stereotype threat. *Applied Developmental Psychology, 24*, 645-662.
- Gliner, J., & Morgan, G. (2000). *Research methods in applied settings: An integrated approach to design and analysis*. Mahwah, NJ: Lawrence Erlbaum.
- Greenleaf, R. (1995). Servant leadership. In J. T. Wren (Ed.), *Leader's companion* (pp. 18-24). New York: Free Press.
- Grow Your Own Teachers: An Illinois Initiative* (n.d.). Retrieved August 17, 2008, from <http://www.growyourownteachers.org/>.
- Gulek, C. (2002). Preparing for high stakes testing. *Theory in Practice, 42*(1), 42-50.
- Hamilton, L., Marsh, J., Mcombs, J., Robyn, A., & Stecher, B. (2007). *Standards-based accountability under No Child Left Behind: Experiences of teachers and administrators in three states*. Santa Monica: Rand Corporation.
- Handlin, O. (1959). *John Dewey's challenge to education: Historical perspectives on the cultural context (The John Dewey society lectureship - number two)*. New York: Harper & Brothers.
- Hoover, R. L. (2003). Ohio proficiency test information, sources, and movements. Retrieved March 23, 2005, from <http://cc.yzu.edu/~rlhoover/OPTISM/index.html>
- Introduction to the Texas accountability system (2002). Retrieved March 22, 2005, from <http://www.tea.state.tx.us/perfreport/account/2002/manual/sec01.pdf>
- Josselson, R., Lieblich, A., & McAdams, D. P. (Eds.). (2003). *Up close and personal: The teaching and learning of narrative research*. Washington DC: American Psychology Association.
- John McCain on Education*. (n.d.). Retrieved June 12, 2008, from http://www.education.com/magazine/article/John_McCain/.
- Kennedy, E. & Miller, G. (2002) Kennedy and Miller call on Bush to back education spending increase to ensure success of historic school reform. *Committee on education and the workforce*. Retrieved November 12, 2003, from http://edworkforce.house.gov/democrats/re_131802.html
- Klein, A., & Hoff, D. (2008, June 9). *Candidates are at odds over K-12 but McCain and Obama both back NCLB*. Retrieved June 11, 2008, from www.edweek.org/ew/articles/2008/06/11/41election.h27.html?tmp=610478353.

- Klein, A., & Hoff, D. (2008, June 9). *Candidates are at odds over K-12 but McCain and Obama both back NCLB*. Retrieved June 11, 2008, from www.edweek.org/ew/articles/2008/06/11/41election.h27.html?tmp=610478353.
- Klein, S. P., Hamilton, L.S., McCaffery, D.F. & Stecher, B. M. (2000). What do test scores in Texas tell us? RAND Corporation web site. Retrieved September 25, 2008 from http://www.rand.org/pubs/issue_papers/IP202/index2.html
- Kline, K. (2008). Hardwired to connect: The new scientific case for authoritative communities. In *Authoritative Communities: The Scientific Case for Nurturing the Whole Child. : Search Institute Series on Developmentally Attentive Community & Society*. (pp. 3-).
- Kober, N., Chudowsky, N., & Chudowsky, V. (2008). *Test scores have been on the rise since NCLB was enacted*. Retrieved August 14, 2008 from <http://www.centerforpubliceducation.org>.
- Kornhaber, M. L. (2004). Appropriate and inappropriate forms of testing, assessment and accountability. *Educational Policy*, 18(1), 45-70.
- Kurosawa, A. (Director). (1950). *Rashomon - Criterion Collection* [Motion Picture]. Germany:Criterion.
- Lao-tzu. (1995). Tao te Ching. In J. T. Wren (Ed.), *Leader's companion* (pp. 69-71). New York: The Free Press.
- Lingle, C. (2005, January 20). Kruse principal resigns. *Fort Collins Coloradoan*, p. 1.
- Lips, D. (2005, Dec. 30). *FOXNews.com - Grading 'No Child Left Behind' - Opinion*. Retrieved June 12, 2008, from <http://www.foxnews.com/story/0,2933,180260,00.html>
- Magnet, M. (1999, February 5, 1999). What is compassionate conservatism? *Wall Street Journal*. Manhattan Institute web site. Retrieved February 10, 2008 from http://www.manhattan-institute.org/html/_wsj-what_is_compassionate_con.htm
- Martinez, T. P., & Martinez, A. P. (2002). Texas tragedy: No Hispanic child left behind? *Education Digest*, 68(2), 35-41.
- Matthews, J. (2003, November 11). No child left behind: Facts and fiction. *Washington Post*, p. A08.
- Mills, C. M., & Keil, F. C. (2005). The development of cynicism. *Psychological Science*, 16(5), 385-391.

- Mulvenon, S. W., Connors, J. V., & Lenares, D. (2001). Impact of accountability and school testing on students: Is there evidence of anxiety? (pp. 1-32):Little Rock. Educational Resources Information Center.
- NEA: *Court Hears NEA Appeal in NCLB Unfunded Mandates Lawsuit*. (2006, October 29). Retrieved August 16, 2008, from <http://www.nea.org/newsreleases/2006/nr061129.html>.
- NEA: *NEA's Positive Agenda for the ESEA Reauthorization*. (n.d.). Retrieved August 14, 2008, from <http://www.nea.org/esea/posagendaexecsum.html>.
- Neill, M. (1996, April 9). *How the principle indicators for student assessment systems should affect practice*. Paper presented at the American Educational Research Association Annual Meeting, New York, NY.
- No Child Left Behind*. (n.d.). Retrieved June 4, 2008, from www.sourcewatch.org/index.php?title=No_Child_Left_Behind_Act.
- Olson, L., & Robelen, E. W. (2001). ESEA passage unlikely before fall. *Education Week*, 20(42), 1,38-39.
- Palmer, P., J. (1998). *The courage to teach*. San Francisco: Jossey-Bass.
- Pamphilon, B. (1999). Zoom model. *Qualitative Inquiry*, 5(3), 393-410.
- Pipho, C. (1999). Eagle eyes and rabbit ears: Locating policy issues. *Phi Delta Kappan*, 80(9), 645-647.
- Poetter, T., & Wegwert, J. & Haerr, C. (2006). *No Child Left Behind and the illusion of reform: Critical essays by educators*. Lanham, MD: University Press Of America.
- Posner, D. (2004). What's wrong with teaching to the test? *Phi Delta Kappan*, 85(10), 749-751.
- President signs landmark education bill. (2002). from <http://www.whitehouse.gov/news/releases/2002/01/20020108-1.html>
- Robelen, E. W. (2000). Time grows short on ESEA renewal. *Education Week*, 19(38), 1, 25.
- Robelen, E. W. (2001a). Bush, Democrats compromise as ESEA bills take shape. *Education Week*, 20(36), 27-33.

- Robelen, E. W. (2001b). ESEA conferees reach accords on social issues. *Education Week*, 21(10), 36-38.
- Robelen, E. W. (2001c). House GOP tosses ESEA-reauthorization bill into the ring. *Education Week*, 20(28), 24.
- Robelen, E. W. (2001d). House overwhelmingly passes sweeping education reform bill, *Education Week*, 20(15). Retrieved February 11, 2008, from http://www.edweek.org/ew/ew_printstory.cfm?slug=15seesa_web2.h21
- Robelen, E. W. (2002). An ESEA primer. *Education Week*, 21(16), 28-29.
- Robelen, E. W. (2004). States given more leeway on test rule. *Education Week*, 23(30), 22, 27.
- Sacchetti, M. (n.d.). *State to consider raising MCAS bar*. Retrieved August 15, 2008, from http://massparents.org/news/2005/raising_mcas_bar.htm.
- Sack, J. L. (2000). Impending retirements shape ESEA debate. *Education Week*, 19(19), 24-26.
- School accountability reports. Retrieved September 25, 2008 from <http://reportcard.cde.state.co.us/reportcard/CommandHandler.jsp>
- Seitsinger, A. M., Felner, R. D., Brand, S., Shim, M., & Dumas, T. (2005, April 2005). *Contexts for learning in the middle grades: Assessment for school accountability*. Paper presented at the Annual Meeting of the American Research Association, Montreal, Quebec.
- Senate passes education bill; school research still threatened. (2001). Retrieved March 22, 2005, from <http://www.cossa.org/privacybill.html>
- Simpson, D. (2006). *John Dewey (Peter Lang Primer)*. New York: Peter Lang Publishing.
- Smith, M. (2002.). *Howard Gardner, multiple intelligences and education*. Retrieved January 7, 2009, from <http://www.infed.org/thinkers/gardner.htm>.
- Solorzano, R. (2008). High Stakes Testing: Issues, Implications, and Remedies for English language learners. *Review of Educational Research*, 78, 260-329.
- Some of John Dewey's writings about inquiry, ethics, the individual and community, education and democracy*. (n.d.). Retrieved August 2, 2008, from www.siu.edu/~deweyctr/shortreadinglist.pdf.

- Stalings, D. T. (2002). The U.S. Department of Education, 1979-2002. *Phi Delta Kappan*, 83(9), 677-683.
- Steptoe, S. (2007, May 30). *How Nebraska Leaves No Child Behind - TIME*. Retrieved Sep. 24, 2008, from <http://www.time.com/time/nation/>
- Stossel, S. (2000). Hard lessons. Retrieved November 15, 2003, from <http://www.theatlantic.com/unbound/interviews/ba2000-11-01.htm>
- Sunderman, G. L., Kim, J. S., & Orfield, G. (2005). *NCLB meets school realities*. Thousand Oaks, CA: Corwin Press.
- Tracey, C. (2005). Listening to teachers: Classroom realities and NCLB. In G. L. Sunderman, J. S. Kim and G. Orfield, *NCLB Meets School Realities* (pp. 81-103). Thousand Oaks, CA: Corwin Press.
- Tyack, D., B., & Cuban, L. (1995). *Tinkering toward utopia*. Cambridge, MA: Harvard University Press.
- The U.S. Response to Sputnik. *Why Files* Retrieved March 22, 2005, from <http://whyfiles.org/047sputnik/main2.html>
- Tulley, M., & Lian-Hwang, C. (1995). Student Teachers and Classroom Discipline. *Journal of Educational Research*, 88, 164-171.
- Unger, C. (2003). Why listen to students? Retrieved April 5, 2005, from <http://www.newhorizons.org/voices/unger.htm>
- Wake-up call over teen stress. Retrieved August 4, 2005, from <http://healthorbit.ca/NewsDetail.asp?opt=1&nltid=155140205>
- Washington, B. (n.d.). *NEA: Republican presidential debate fails to focus on 'No Child Left Behind'*. Retrieved June 12, 2008, from <http://www.nea.org/newsreleases/2007/nr070605.html>.
- What Is Compassionate Conservatism?* (n.d.). Retrieved February 10, 2008, from http://www.manhattan-institute.org/html/_wsj-what_is_compassionate_con.htm
- Wixson, K., K., Dutro, E., & Athan, R. G. (2003). The challenge of developing content standards. In R. E. Floden (Ed.), *Review of research in education* (Vol. 27, pp. 69-107). Washington DC: American Educational Research Association.
- Wolcott, H. (1994). *Transforming qualitative data: Description, analysis, and interpretation*. Thousand Oaks, CA: Sage Publications (pp. 21-22)

Wolk, R., A. (1998). Education's high stakes gamble
Education Week, 18(15), 48.

Young, S. (n.d). No child left behind history. Retrieved June 14, 2005 from
<http://www.ncsl.org/programs/educ/NCLBhistory.htm>

EPILOGUE

As a classroom teacher who has been teaching for nearly three decades I have witnessed a great number of changes in education. I have seen reforms come and go, and watched more than a few battles rage over competing instructional methodologies. During the years when the standards movement was gaining momentum I was an enthusiastic participant. I served on a team that examined new science standards and I wrote a standards-based unit for the Colorado State Department of Education. But when our governor pushed forward a plan to assess students with annual high-stakes testing, I became skeptical. The verdict was still out on standards as far as I was concerned but I was sure the testing idea was not a good one.

In the years that followed, my suspicions about testing and reporting scores were confirmed. Schools in my district with high minority or low-income student populations experienced declines in numbers as affluent parents pulled their children to send them to other district schools with higher test scores. Some of these schools were charter schools.. These schools posted excellent scores and filled quickly with homogenous populations of mostly affluent White students.

I observed changes in my classroom as well. I watched my students as they took their annual CSAP exams. Most were compliant, but some were obviously in distress. For a number of years children identified with emotional disorders were included in my classroom. Some of these students made their frustrations with testing known in dramatic ways including loud protests, crying, and, in one instance, by overturning a desk.

I felt pressure too. Pressure to make my students perform well on the tests. I taught my students testing strategies so they would know how to approach standardized tests. I taught them skills I knew were likely to be on the test and I taught skills I thought were not developmentally appropriate or necessary in order to prepare them for the exams.

All this made me angry. I expressed this anger in various ways. I joined with colleagues and carped about the testing program in the teacher's lounge. I posted a faux news release satirizing the testing program on the staff bulletin board. I commiserated with parents who were critical of the tests. But these outlets were not enough. I was still angry. I could not believe I had allowed myself to become a part of such a flawed plan.

Soon after I enrolled in the doctoral program at CSU I realized that I could do something with my anger. I could direct it towards a study of how assessment affected students. As I took action my anger began to subside. I still am frustrated about high-stakes testing and I am frustrated by its impact upon the quality of education and about how testing impacts students. But this research process has allowed me to feel I have taken some action against what I believe to be flawed policy.

Conducting this study has helped me in another way. It has given me hope. I have come to realize there are many others who think the current accountability system is poorly conceived and implemented. Legislators from both political parties are publicly declaring their dissatisfaction with No Child Left Behind. Parents are unhappy with the system. School board members are upset about unfunded mandates. It is not just cranky teachers and academics that question the process. I think we are reaching a critical level of sentiment directed against the current system. I think we will see changes in the next

authorization of the Elementary and Secondary Education Act. The changes are not likely to be as radical as I would like to see but I think it will be changed in ways that will lessen its harmful effect upon students.