

12-2009

# Teachers' Perceptions of the Effects of the AIMS Test on Arizona High School Math and English Curriculum and Instruction

Joseph Leland Heywood  
*Utah State University*

Follow this and additional works at: <http://digitalcommons.usu.edu/etd>

 Part of the [Curriculum and Instruction Commons](#)

---

## Recommended Citation

Heywood, Joseph Leland, "Teachers' Perceptions of the Effects of the AIMS Test on Arizona High School Math and English Curriculum and Instruction" (2009). *All Graduate Theses and Dissertations*. Paper 490.

This Dissertation is brought to you for free and open access by the Graduate Studies at DigitalCommons@USU. It has been accepted for inclusion in All Graduate Theses and Dissertations by an authorized administrator of DigitalCommons@USU. For more information, please contact [dylan.burns@usu.edu](mailto:dylan.burns@usu.edu).



TEACHERS' PERCEPTIONS OF THE EFFECTS OF THE ARIZONA INSTRUMENT  
TO MEASURE STANDARDS (AIMS) TEST ON ARIZONA HIGH SCHOOL  
MATH AND ENGLISH CURRICULUM AND INSTRUCTION

by

Joseph L. Heywood

A dissertation submitted in partial fulfillment  
of the requirements for the degree

of

DOCTOR OF EDUCATION

in

Education

Approved:

---

Martha Whitaker, Ph.D.  
Major Professor

---

Sylvia Read, Ph.D.  
Committee Member

---

Ronda Menlove, Ph.D.  
Committee Member

---

Francine Johnson, Ph.D.  
Committee Member

---

Steven Camicia, Ph.D.  
Committee Member

---

Byron R. Burnham, Ed.D.  
Dean of Graduate Studies

UTAH STATE UNIVERSITY  
Logan, UT

2009

Copyright © Joseph L. Heywood, 2009

All Rights Reserved

## ABSTRACT

Teachers' Perceptions of the Effects of the Arizona Instrument to Measure  
Standards (AIMS) Test on Arizona High School Math and  
English Curriculum and Instruction

by

Joseph L. Heywood, Doctor of Education

Utah State University, 2009

Major Professor: Martha Whitaker, Ph.D.  
Department: Elementary Education

This is a doctoral dissertation centered on a qualitative research case study on teachers' perceptions of the effects of a particular high-stakes test—the Arizona Instrument to Measure Standards (AIMS) on high school English and math curriculum and instruction. Specifically, this case study involved three focus group interviews at three Arizona high schools. Participants were experienced English and math teachers at each school site. To broaden the perspective gained from the focus group interviews, the study included classroom observations and document analysis. Two of the high schools chosen are located on the Navajo Reservation and were specifically chosen to include an important population of students in a discussion in which they are largely forgotten even though they face unique challenges. A review of the literature on high-stakes testing effects reveals almost no extant literature on the effects of the AIMS test. This study will

contribute to the national dialogue on testing effects as well as make an important foundational research contribution to the Arizona state dialogue on AIMS effects.

(149 pages)

## DEDICATION

For Dr. Harold L. Heywood, who showed me that I could.

## ACKNOWLEDGMENTS

I would like to thank my dissertation committee chair, Dr. Martha Whitaker, for her guidance and direction throughout the entire dissertation process. Without her important contributions, I would not have completed this project. I would also like to thank my committee members, Dr. Ronda Menlove, Dr. Sylvia Read, Dr. Steven Camicia, and Dr. Francine Johnson, for their suggestions and direction as we worked to make this dissertation a success. I also recognize the unseen work of Dr. Deborah Byrnes whose energy and tireless work have made this doctoral program possible. Finally, I thank my wife and best friend, Brandi Heywood, for the many hours of sacrifice on my behalf as I worked to complete this project.

Joseph L. Heywood

## CONTENTS

	Page
ABSTRACT.....	iii
DEDICATION.....	v
ACKNOWLEDGMENTS .....	vi
LIST OF TABLES.....	ix
CHAPTER	
I. INTRODUCTION .....	1
Theoretical Framework.....	4
A Brief History of Standardized Testing .....	5
II. LITERATURE REVIEW .....	15
John Dewey, Progressivism, and Democracy.....	15
Standardized Testing and High-Stakes Testing .....	18
Philosophical Arguments on Testing Issues .....	19
Test Score Analysis Studies.....	22
Testing Effects Research: 1990-1999 .....	25
Testing Effects Research: 2000-2008 .....	27
AIMS Test Research .....	39
Conclusion .....	40
III. METHODS .....	42
Theoretical Foundations.....	42
Case Study .....	46
Focus Group Interviews.....	51
Participant Selection .....	52
Data Collection Protocol.....	54
Credibility and Trustworthiness of Data.....	55
Limitations and the Role of the Researcher.....	57
IV. RESULTS .....	60
Participant School Demographics.....	60



Page

Clarification of Important Terms .....	62
Emergent Themes in the Data.....	65
Classroom Observations and Document Collection .....	108
V. ANALYSIS, INTERPRETATION, AND IMPLICATIONS .....	114
Observations .....	114
Conclusion .....	125
REFERENCES .....	130
VITA.....	138

## LIST OF TABLES

Table	Page
1. Participant Schools' AIMS Scores and AYP Status .....	61
2. Participant Schools' Demographic Information .....	61
3. Participant Schools' 2008 Ethnicity Percentages .....	63

## CHAPTER I

### INTRODUCTION

In the year 2000, Arizona legislation (Education 2000/Proposition 301) gave the Arizona Department of Education (ADE), among other things, funding and authority to develop a statewide system of school accountability. In 2002, the state legislature of Arizona passed A.R.S. §15-241 (or *Arizona Learns*) which established a research-based evaluation system by creating achievement profiles and designated classifications (excelling, improving, maintaining, underperforming, and failing to meet academic standards) for schools. Among other data, the achievement profiles and classifications are based on student AIMS (Arizona's Instrument to Measure Standards) test scores.

The initial Education 2000/Proposition 301 legislation has been supplemented or amended with subsequent state and federal legislation. For example, House Bill 2277 amended A.R.S. §15-241 in 2003 and, most importantly, the federal No Child Left Behind (NCLB) legislation in 2002 strongly influenced the following important state stipulations: schools must assess 95% of the student population using AIMS, and each school must meet the state's proficiency targets (which NCLB requires to escalate to 100% of students achieving proficiency by the 2013-2014 school year).

With the AIMS test playing such an important role in school classifications (which in turn affect funding and programs), the test also affects individual students because passing the AIMS is required for graduation. All high school students take the high school AIMS test in the spring of their sophomore year for the first time. They can pass or fail three sections (math, reading, and writing). If a section is passed, the student

no longer takes that portion of the exam. If a student fails a section, then he can take the exam again in the fall and spring of his junior and senior years. Every student must pass the exam to graduate (though there are some exceptions made for special education and other students, the general population must pass the exam). Though the AIMS test only covers reading, writing, and math, a science test (not required for graduation) was introduced the 2007-2008 school year. As of the 2008-2009 school year, the Math, Reading, and Writing sections of the AIMS test are the only high-stakes (required for graduation) portions of the exam. Finally, any student achieving an Exceeds on all three sections of the AIMS qualifies (if certain GPA criteria are met) for scholarships to any state college or state university.

Though the state does provide test scores to districts with instructions that districts provide these to teachers for analysis and, potentially, adjusted practice, how curriculum and instruction are affected by these test scores is largely unknown. Analysis of test scores will reveal little more than a rise or fall in scores, and one might say, “Hey, test scores are on the rise, teaching is getting better, and students are learning at higher levels.” The following literature review includes a history of standardized testing and a review of the extant literature on high-stakes testing effects. The history of standardized testing, and this review of the literature suggest a need for more research on the effects of high-stakes testing, particularly regarding the effects on the actual classroom curriculum and instruction.

An important aspect of this study is the inclusion of two Native American high schools located on the Navajo Reservation in northeastern Arizona. Thousands of Navajo

students attend state schools located on the Navajo Reservation. These schools are part of the ADE and must follow all state rules and guidelines regarding public education, including graduation requirements associated with the AIMS testing. These students face significant challenges regarding poverty, exposure to English as a native language, and other important factors. Yet, these Navajo students are a largely forgotten part of the discussion on AIMS testing and AIMS effects. It is my purpose here to include these two schools in this study to intentionally position them prominently within the discussion on AIMS testing so that the voices of the educators servicing these Native American students might be heard.

With the power of high-stakes testing—more and more states are creating graduation exams (like Arizona), and high-stakes tests stand as gatekeepers to our systems of higher education—it is remarkable that so many researchers are suggesting that the body of research on high-stakes testing effects is somewhat shallow. The effort to visit the classrooms and talk to the teachers about the possible effects of these high-stakes exams has been relatively meager. The need for more researchers to go into public schools and classrooms and ask important questions about testing effects is clear. The purpose of this study, the questions I have sought to answer, are centered on finding out how the AIMS test might be changing curriculum and instruction in Arizona. In particular, what can teachers help us to understand about the way the AIMS test is changing curriculum and instruction? There is currently no published literature seeking to answer these specific questions.

## Theoretical Framework

The theoretical foundations for this study are grounded in philosophy of Dewey (1915, 1916, 1938). Most compelling to me are Dewey's ideas of democratic progressivism and its tenets of child-centeredness, critical inquiry, open and dynamic systems, active participation, cultural diversity, and an appreciation of the artistic and creative in children and education.

Dewey's views inform the philosophical contradiction between what many perceive to be valuable educational experiences and standardized assessment. Dewey (1938) focused his life's work on the democratic education of the individual child. He broke with traditional, "old" views of education focused on memorization, structure, and what we might refer to today as "standardization." He promoted the development of curriculum which focused on experience based on an individual child's needs, tying together both what she knew with what she needed to know to be part of a larger society. Dewey insisted that inquiry and problem solving, reflection and critical thinking, should all be part of a broader curriculum supplanting the packaged, rote learning in traditional schools.

My theoretical lens assumes the goodness of Dewey's democratic progressivism. My research questions were driven by my own beliefs that align with Dewey's democratic progressivism. These beliefs and my concerns about the mismatch between democratic progressivism and current educational policy and practices informed every phase of my study.

The methods of inquiry for this study were grounded in Eisner's principles of

educational criticism (which will be more fully summarized in Chapter III); when applied to an evaluation of an aspect of education, Eisner's educational criticism can broaden perspectives about the status of our current school system and the effects of high-stakes testing practices. Eisner believed in revealing the intricacies of education through a qualitative approach to research focused on four aspects of criticism: descriptive, interpretive, evaluative, and thematic. The research narratives created by following his educational criticism principles can help us to better understand the pedagogy of schools and the dynamics of the classroom experiences of teachers and students.

### A Brief History of Standardized Testing

To gain perspective on how our current education system largely left behind Dewey's ideas about our schools being child centered, a review of the rise of standardized testing is needed. Standardized testing was suggested and developed in scattered, and usually isolated, parts of the developing educational system in America throughout the 1800s. However, no wide-scale testing efforts developed until the 20<sup>th</sup> century. As the U.S. moved into the 1900s, school leaders began restructuring public schools around the ideas of meritocracy—the belief that students who demonstrated ability merited more opportunities. As the public school system developed in the first half of the 20<sup>th</sup> century, school leaders, psychologists, and test developers began working in concert to bring about a structured and standardized school system. Much of the early blending of public schools and test development was founded on the earliest versions of our modern multiple-choice standardized exams—the IQ test (Tyack & Hansot, 1982).

Building on the work of French psychologist, Alfred Binet, Terman (1916 as cited in Hanson, 1993) created an IQ test at Stanford University (the Stanford-Binet). Binet's intention had been to identify mental deficiencies for special education services; Terman had a larger vision for IQ tests: he "held that intelligence testing would facilitate the placement of people in those educational programs and vocations for which their endowments best suit them" (Hanson, p. 210). This idea was to become a foundational principal for the flourishing of standardized testing throughout the American public education system over the next 90 plus years.

Seeing testing as a means to reveal innate qualities that could be used to demonstrate merit, the U.S. government began administering IQ tests to screen military applicants as early as World War I and to immigrants freshly arriving in America in the early decades of the 1900s (Hanson, 1993; Kamin, 1974; Sacks, 1999). The tests identified so-called strengths and weaknesses in a person's intelligence, and the results were used to sort people into categories of service. For decades, IQ tests were seen as a powerful tool for revealing a person's innate abilities; even today, Americans recognize the stigmas, both positive and negative, associated with having a "high IQ" or a "low IQ."

By 1926, the Standardized Achievement Test (SAT) had been developed as a means of screening college applicants, and by the 1940s, it had replaced the essay as the favored and primary method of screening college bound students. Hanson (1993) suggested this transition was due to the explosion of secondary and higher education student populations; "In 1870, about 80,000 students attended some 500 secondary



schools. By 1910, the number of secondary school students had grown to 900,000” (p. 213). Moreover, the growth continued. “In 1910...19% of 15- to 18-year-olds were enrolled in a public or private high school. By 1940...73% of American youth were enrolled in high school” (Goldin & Katz, 2008, p. 195). Goldin and Katz suggested this growth was due in part to increased enforcement of states’ child labor and compulsory schooling laws. At the same time secondary populations were booming, higher education experienced a significant increase in student numbers, growing at 5 times the rate of the population between 1890-1924 (Goldin & Katz; Hanson). With so many new students in public education, educators and government leaders began searching for ways to measure the schooling masses in an efficient manner.

In the 1940s, the Educational Testing Service (ETS) was founded. Over the next 60 years, ETS would come to dominate the American testing landscape by developing and overseeing a wide variety of tests: many state tests for public education K-12, the SAT for undergraduate college admissions, the graduate records examination (GRE) for graduate school admissions, the advanced placement (AP) exams for advanced high school students, and many significant tests for school systems in other nations. The ETS has played a central role in the development of American standardized testing and our almost complete reliance upon standardized tests to evaluate American children and college adults in our school systems.

As graduation from American high schools became an expected norm rather than a novelty (29% national graduation rate in 1920 compared to the at least “expected” graduation of all students by the 1960s), this provided impetus for the standardized

testing movement (Mitchell, 1992). Mitchell also suggested that the development and widespread use of technology in the 1960s and 1970s allowed the rapid development of the testing industry; “As computers, scantrons, and similar machines became cheaper and easier to use, norm-referenced, multiple-choice testing became almost universal” (p. 172). Concurrent with new expectations for graduation and advances in technology, some argue that school administrator power and respect, gained in the first half of the 20<sup>th</sup> century (Tyack & Hansot, 1982), began to wane as reformers, citizens, and politicians began distrusting public school leadership and calling for more accountability and change and shifting power over schools to both the state and federal governments (Dorn, 2007).

The 1960s were a turning point in American public schools. Prior to this decade, U.S. public education was a relatively disjointed body of districts, state education agencies, and some federal programs. The power in education was primarily localized and the federal government played a limited role in education.

In the United States education governance is not a power constitutionally granted to the federal government. Because of this, education policy making has historically been situated primarily at the local level. Until the recent reforms, curriculum, teaching, and assessment decisions were made at the local school and district (town, city or county) levels. State education departments generally ensured that all schools met minimal requirements regarding graduation, the length of the school day and year, physical facilities, and teacher preparation. (Hursh, 2005, p. 606)

As the shifting of power gave federal and state governments more control, the standardization of schools led to educational policy that “intended to divorce the public from the governance of public schools” (McNeil, 2000, p. 230). Beginning with important legislative changes at the federal level in the 1960s, “new state and federal regulations have transferred power away from teachers, parents and local community

members and towards corporate and political leaders at the state and federal levels”

(Hursh, 2005, pp. 605-606).

The common regime of low-stakes, diagnostic, norm-referenced achievement testing began to change, initially very slowly, in the 1960s. Until then, the use of test scores to monitor the performance of education systems was for the most part limited to local districts that chose to monitor the performance of their own schools. In the 1960s, two actions by the federal government began to change this the establishment of the ESEA and the NAEP]. (Koretz, 2008, p. 54)

In 1965, the Elementary and Secondary Education Act (ESEA) formalized federal spending on education and outlined federal regulations for public education receiving these funds. “The ESEA of 1965 [represents] the beginning of a dramatically expanded federal role in K-12 education” (Manna, 2007). Koretz (2008) wrote,

The ESEA of 1965...established the Title I compensatory education program—the precursor of today’s No Child Left Behind—to improve the performance of students in low-income schools. This marked the first major involvement of the federal government in funding and directing general elementary and secondary education. The law also mandated evaluation of the Title I program (the first time that federal legislation establishing a major social program required a formal program evaluation). In 1974 Congress established the Title I Evaluation and Reporting System (TIERS), which based required evaluations of Title I programs on students’ scores on standardized, norm-referenced achievement tests. (pp. 54-55)

Since 1965, the ESEA has been reauthorized repeatedly (with significant changes) under different names. In 1970, the federal government began the National Assessment of Educational Progress (NAEP), a national test used to determine regional trends in academic performance, which influence federal policy. In 1979, the Secretary of the federal Department of Education became a presidential cabinet position. By 1980, school accountability had become a weighty issue for the federal government.

Standardized testing grew throughout the 1980s and 1990s as the federal

government placed more and more emphasis on its merits. With the release of, *A Nation at Risk*, the National Commission on Excellence in Education (NCEE, 1983), a commission created by Secretary of Education, Terrell Bell, reported American public school standards as too low, and among other recommendations, called for greater accountability in public schools; a great reform movement commenced in which standardized testing and accountability shifted into prominent roles. In 1988, Congress created the National Assessment Governing Board, whose members are appointed by the Secretary of Education and oversee the NAEP. The same year was marked by important changes to the ESEA:

The ESEA reauthorization of 1988 (P.L. 100-297) began to reflect the increasing importance of states as educational reformers.... Not only did the states appear explicitly in the [Title I] statement of policy, but the law also began to link academic performance of Title I students to state-defined achievement levels as a way to identify poorly performing schools. (Manna, 2007, p. 73)

With each new ESEA reauthorization, the federal role in educational reform and public school evaluation increased dramatically and standardized testing was at the center of this accountability reform movement.

In 1989, newly elected President George H. W. Bush convened an education summit at the University of Virginia where, for the first time in U.S. history on a matter of education, the President met with, among others, governors from all 50 states. The outcome of this meeting was extraordinary:

For the first time in the nation's history, the chief executives pledged to set national goals for education and to hold themselves accountable for attaining them. The goals, moreover, were ambitious; some called them unrealistic. Announced in early 1990, they stated that by 2000, every child will enter school ready to learn; the high school graduation rate will increase to at least 90 percent; all students will demonstrate competence in challenging subject matter; U.S.

students will be first in mathematics and science achievement; all adults will be literate; and all schools will be safe and drug-free. (Rothman, 1995, pp. 111-112)

Of course, all of this accountability required measurement, “officials responding to the goals began to call for national standards for student performance and a related national system of assessments” (Rothman, p. 112).

With the federal government firmly ensconced in the world of education, leaders and policy-makers from both major political parties at both the federal and state levels began pushing national and state standards and the creation of assessments to monitor student achievement. With both major political parties on board at both the national and state levels, standardized testing swept through all aspects of public education relatively unchecked, at least politically, through the 1990s and into the new millennium (McDonnell, 2004). With back-to-back presidents, Bill Clinton and George W. Bush, from 1992-2008, standardized testing was pushed from the desk of the U.S. president like never before. In 2002, President Bush signed the NCLB legislation (another reauthorization of the ESEA with powerful changes) which now required the testing of all public school students at specified grade levels each year and further solidified the federal government’s stance that public education will have clear performance standards and measurable objectives. The consequences of NCLB include the creation of state tests which are used to rate children, determine graduation from high school, rate schools, rate teacher performance, determine teacher licensure and performance, and in some cases determine teacher pay; though it is important to note that these consequences of standardized testing were being identified even before NCLB (for example, see Smith, 2000, p. 336); NCLB amplified many of these consequences and made them consistent

throughout all 50 states.

The shifting of power from schools and communities to the federal and state governments has restructured our schools academically, bureaucratically, and financially. “In 1967, 80 percent of a school’s budget was devoted to regular instruction. By the late 1990s, the percentage of funds devoted to regular instruction had dwindled to about 50 percent” (Baines & Stanley, 2004, p. 8). Baines and Stanley claimed that the 2003 annual cost of high-stakes testing was “somewhere between \$20 and \$50 billion, or 5.5% to 14% of every dollar spent for public schools” (p. 8). An interesting note on the shift of power from local schools to the federal government involves the breakdown of funding for public schools: “In America, public school funding is available from three main sources, the federal government, who typically contribute about 7-8%, the state (contributing around 48%) and local taxes (around 45%), although the exact proportions do vary by state” (Smith, 2004, p. 514). One of the arguments for federal “takeover” of public education, despite the fact that the federal government contributes such a small percentage of the public school funding, is the discrepancy between poor communities (low property values mean less money collected for schools and thus lower quality schools) and wealthier communities (higher property values generate greater income for the “rich” schools). The federal government argues that their standardization and accountability movement equalizes opportunity and forces underperforming schools (typically in poor communities) to find ways (often through special federally funded programs) to raise student achievement, which is monitored of course through standardized exams.

Two years after NCLB was signed into law, Goertz and Duffy (2003), identified four major challenges facing schools and districts as they implement NCLB. First, state assessment programs will need to expand considerably. This expansion will come at a major cost (only partially aided by federal money) and will require some states to drop assessment of subjects like social studies and “other subjects not covered by the NCLB Act. This could have the unintended consequence of narrowing educators’ focus on the tested subjects of reading and mathematics” (p. 9). Many districts have chosen to eliminate local assessments, closely connected to curriculum, to save money and time for administering state exams for NCLB. Second, educators are concerned about whether a single state exam can serve multiple purposes.

Policy makers expect one assessment system to provide indicators of the performance of the education system, hold schools and educators accountable for their performance, certify student performance as students move from grade to grade or out of the K-12 education system, motivate students to perform better and teachers to change their instructional content and strategies, and aid in instructional decisions about individual students. (Goertz & Duffy, p. 9)

Third, as educators respond to the NCLB threat of sanctions due to underperformance on state exams, “the question then becomes, will educators pay attention to the right kinds of student performance data, which is a function of the quality and appropriateness of the test, and will they know how to act on that data?” (p. 10). Finally, incentives and goals don’t improve instruction and student performance; true change requires “meaningful professional development” and “capacity building” (p. 10). Improving instruction requires a vast amount of resources, financial and human, to meet the high standards of NCLB. The authors suggest our current system does not have the capacity (especially financially) to provide these resources of change.

With this proliferation of standardized testing and the prominent and controlling role the federal government and state governments have come to play in these national and state testing systems, the body of research on the effects of standardized testing has likewise begun to expand. Researchers and educators are asking tough questions about how these standardized tests are influencing our public school system and the children the system serves. This study will contribute to this effort.

This study focused on the question: What effects is the high-stakes AIMS test having on high school English and math curriculum and instruction in Arizona. Three focus group interviews were conducted—one interview per site at three different Arizona high schools. Teachers were prompted with questions about the AIMS test throughout the interviews, but the discussions were generally free to move in the directions the teachers involved decided to go. The interviews were recorded, transcribed, and analyzed looking for general themes or important ideas emerging out of the three groups. Classroom observations and document analyses further contextualized the interview data. I also examined the similarities and differences between the schools and considered the way these institutional dynamics may have contributed to these themes. I sought to keep the teachers' voices in the study by providing many direct quotes throughout the data section of this research report. I wanted the readers of this study to hear the teachers as directly as possible. The study concludes with an analysis of the interviews and subsequent classroom observations and document analysis. In closing, I offer some conclusions as a researcher based on my interactions with these teachers at their schools and their feelings and ideas about the AIMS.



## CHAPTER II

### LITERATURE REVIEW

This literature review builds a foundation for the research conducted in this study. First, I will summarize the basic principles of Dewey's democratic progressivism. Next, I will turn to the review of standardized testing. I will begin by clarifying the use of the terms "standardized testing" and "high-stakes testing" in this study. Third, some of the philosophical arguments against standardized testing will be noted as they were the impetus for this study as I wondered about the veracity of their claims. Fourth, I will briefly look at some of the examples of quantitative test score analysis studies as they are an important genre in the larger research base on testing effects. Fifth, the testing effects studies of the 1990s, the decade when testing effects studies really began to develop with frequency, will be reviewed. Sixth, a large part of the literature review will focus on the most recent research (2000 to present), with a summary of the important studies which examined testing effects by seeking out teachers' and students' perceptions. Finally, I will briefly review the extant published literature on AIMS effects.

#### John Dewey, Progressivism, and Democracy

In breaking with traditional views of schools, Dewey's (1915) progressivism contradicted "the typical points of the old education: its passivity of attitude, its mechanical massing of children, its uniformity of curriculum and method" (p. 23). These words, written in the early 20<sup>th</sup> century, might describe much of what we call public education today in a school system highly dependent upon uniform curriculum and

standardized testing. “The specter of testing has haunted public education since the time of Dewey and Thorndike, giving rise over time to a science of education and ongoing uneasiness about its influence on educational practice” (Gunzenhauser, 2006, p. 241). Modern test scores are often used to classify schools and categorize children into groups (“making AYP,” “high performing,” “failing,” “marginal,” etc.). In 1916, Dewey wrote the following, which again could easily describe many of the test-driven school systems today.

The vice of externally imposed ends has deep roots. Teachers receive them from superior authorities; these authorities accept them from what is current in the community. The teachers impose them upon children. As a first consequence, the intelligence of the teacher is not free; it is confined to receiving the aims laid down from above. Too rarely is the individual teacher so free from the dictation of authoritative supervisor, textbook on methods, prescribed course of study, etc., that he can let his mind come to close quarters with the pupil's mind and the subject matter. This distrust of the teacher's experience is then reflected in lack of confidence in the responses of pupils. The latter receive their aims through a double or treble external imposition, and are constantly confused by the conflict between the aims which are natural to their own experience at the time and those in which they are taught to acquiesce. Until the democratic criterion of the intrinsic significance of every growing experience is recognized, we shall be intellectually confused by the demand for adaptation to external aims. (p. 81)

Of course, the external aims of Dewey's day were not the thoroughly test-centric aims of the American public education system of today, but Dewey's ideas are equally applicable to our modern system as they were for him in cautioning educators about standardized curriculum and standardized testing which was beginning to develop (especially in the form of IQ testing) in Dewey's time. Placing Dewey in a socio-historical context, his philosophy developed in a time of increasing pressure to control schools due to massive immigration and urban chaos, I.Q. testing, and an increasing standardization of certain aspects of education.

Dewey's progressivism advocated a truly democratic society with education playing a central role in creating and sustaining that democracy. Dewey's definition of democracy centered on the individual in relationship to society; our schools must free the individual to participate in a free society. He wrote,

A society which makes provision for participation in its good of all its members on equal terms and which secures flexible readjustment of its institutions through interaction of the different forms of associated life is in so far democratic. Such a society must have a type of education which gives individuals a personal interest in social relationships and control, and the habits of mind which secure social changes without introducing disorder. (1916, p. 74)

While some authors use this vision of democracy to empower systems of education that embrace high-stakes testing (Raphael & Au, 2005), Dewey used this definition to empower the individual and not a standardized system. Dewey's vision of democratic progressivism developed as others espoused principles of social efficiency progressivism (like Charles W. Eliot), a line of thinking with parallels to philosophy underlying the standardized education principles of today.

For Dewey, the individual was the focus of education. Dewey (1916) implored educators "to find a deeper sense of the function of education in discovering and developing personal capacities" (p. 67). In accordance with Plato's ideas about individuality, Dewey promoted the education of individuals and not classes; yet, he broke with Plato's limited classifications of individuals and asserted that human beings' "original capacities are indefinitely numerous and variable" (p. 68). Dewey's views of the individual called for an educational system filled with variety and flexibility and real experience that each student could connect with on many different levels.

Dewey urged teachers to educate these individuals with a curriculum and with

methods dependent on what we might call “authentic” learning—real experiences for discovering truths and knowledge. Separating a child from reality for the duration of a school day and presenting knowledge in isolation of real experience (pouring facts into an empty vessel) was contrary to the “doing” required in learning. Dewey (1915) wrote,

It is our present education which is highly specialized, one-sided, and narrow. It is an education dominated almost entirely by the mediaeval conception of learning. It is something which appeals for the most part simply to the intellectual aspect of our natures, our desire to learn, to accumulate information, and to get control of the symbols of learning; not to our impulses and tendencies to make, to do, to create, to produce, whether in the form of utility or art. (p. 19)

He felt that our focus on abstract ideas, often disconnected from the day-to-day realities a child faces, does not meet the needs of individual children. As Dewey (1915) said, “The child has not much instinct for abstract inquiry” (p. 30).

Dewey’s progressivism sits in stark contrast to the current state of the American public school system. Dewey’s progressive education is open, dynamic, child-centered, founded on experimental, inquiry-based learning, respectful of diversity, creative, and intended to provide an educational experience rich with critical, investigative inquiry that brings about democracy. The closed, institutionalized, standardized, productivity-based, business and factory model, culturally uniform, test-centered public school system today has elements of Dewey’s progressivism but has largely lost sight of his theoretical principles as policymakers and politicians have slowly taken over public education.

### Standardized Testing and High-Stakes Testing

Much of the literature on testing seems to use the terms “standardized testing” and “high-stakes testing” interchangeably because high-stakes tests are often a subset of

standardized tests. It is important to define the two terms. A standardized test is a uniform test given to a large number of students with scores used for measurement of student progress, a variety of comparisons, and curriculum, school, or teacher evaluations. The term “standardized test” is generally used to refer to both high-stakes tests (tests with consequences) and low-stakes tests (tests without specific consequences). A high-stakes test might refer to a test that is required for graduation (the AIMS test), used to determine a school’s academic standing with the state (the AIMS test), or used to determine college eligibility or placement (and possibly scholarships) for individual students (the AIMS test, the national ACT or SAT). A high-stakes test is one with “serious consequences for students, their teachers, and their educational institutions” (Davis, 2006, p. 1). With the passing of NCLB in 2002, high-stakes testing “now pervades educational practice” (Gunzenhauser, 2006, p. 241). The purpose of this study was to focus on a particular high-stakes test; however, in discussing testing, it is difficult to separate high-stakes testing out from the general literature on standardized testing, especially in reviewing the history and literature of standardized testing. Researchers have not yet made a concentrated effort to create a body of research and literature wholly devoted to the specific phenomenon of “high-stakes testing.”

### Philosophical Arguments on Testing Issues

For the past 20 years, as standardized testing—and especially high-stakes testing—has become increasingly pervasive in public education, the national dialogue on standardized testing has included a variety of philosophical indictments of the high-stakes

testing trend. Stoskopf (2000) has suggested that high-stakes testing is rooted in the eugenics movements of the early 1900s and racist assumptions about IQ. Neill (2003) wrote, “an overemphasis on testing will undermine, not strengthen, the ability of schools to ensure high-quality academic and social experience for all their students” (p. 19). Sirotnik (2002) has written about the increasingly “punitive nature” of our high-stakes practices and is concerned with the increased “dropout rates of marginalized students” (p. 662). Sacks (2000) defined marginalized students in stating that “the losers in high-stakes testing schemes always have been children of the poor, the working class and under-educated. And the winners always have been children of the privileged, well-educated and the affluent” (p. 6). Popham (2000) took to task what he claimed to be the false assumptions underlying the testing culture: “The score-boosting game rests on the assumption that a set of high scores by students equals more successful instruction by educators. That assumption, however, is misguided because it clearly misapplies the information that can be gleaned from standardized tests” (p. 12). Elsewhere, Popham (1999) claimed that the “large-scale-assessment community is unconcerned” with the instruction of children, focusing all of its concern on accountability instead (p. 13).

There is also the claim that the theoretical foundations of standardized testing are rooted in early 20<sup>th</sup> century psychometric theories that no longer hold weight in current educational curriculum or instruction (Gipps, 1994); and the claim that our current stated theories of assessment do not correspond to our current theories of instruction and learning (Willis, 1993). Weaver (2004) wrote, “As the high-stakes paradigm becomes a focus for teaching content as opposed to teaching children, the culture of teaching will be

drastically affected” (p. 258). Gunzenhauser (2006) addressed “the long-term implications of high-stakes accountability on philosophies of education” by analyzing “high-stakes accountability as a problem of a particular kind: the foreclosure on possibilities for our aims for what it means to be an educated subject” (p. 242). Finally, Dantley (2003) suggested the incorporation of a spiritual dimension to school leadership to guide reform efforts in moving schools beyond high-stakes testing: “When educational leadership is ensconced in a spiritual context, it serves to bring about reform and reconstruction that are external to the normative paradigms of school change. Instead of using empirical or quantifiable qualifiers of school effectiveness, educational achievement takes on a more axiological, moral, and teleological tenor” (p. 282).

Most of these writers approach their critique of testing by outlining their philosophical or psychological foundations of learning, and then comparing standardized testing to their foundation, pointing out the disconnect as a failure of the tests (for example, see Eisner, 1982, Chapter 4; 1998a, Chapter 11). It is this theorizing that often triggers both qualitative and quantitative research studies to verify claims; the philosophical writings about high-stakes testing do not overrule scientific inquiry (Curren, 2006); they stimulate research. Thus, as this study was born out of a desire to investigate certain claims philosophers were making about testing effects, it was important to recognize here the contributions of these critics and philosophers to high-stakes testing research.

## Test Score Analysis Studies

One of the key topics of studies researching the effects of standardized exams has been comparative analysis of exam scores. Though this study is patterned after the qualitative inquiry into teachers' and students' perceptions (and thus my literature review will focus on these types of studies), it is important to note examples of test score analysis studies because these are the studies most often used to promote the testing movement. In addition, it is most often these quantitative analyses of test scores that we read about in the mainstream media.

Amrein and Berliner (2002) examined the performance scores of the NAEP exam for 18 states with high-stakes tests looking for effects of the high stakes exam. The authors claimed that what students learned either was not significantly affected or went down as a result of high-stakes testing, and they suggested that serious discussion about high-stakes testing and its purposes is needed. A rebuttal to this study's findings used an "improved" statistical analysis to contradict Amrein and Berliner's findings, criticizing their research and calling for better evaluation of studies claiming to be scientific research on testing effects (Raymond & Hanushek, 2003). In another study, Greene, Winters, and Forster (2003), compared high-stakes tests with scores on other tests and concluded that the high correlation could be interpreted to mean that "when a state's high stakes test scores go up, we should have confidence that this represents real improvements in student learning" (summary page).

Other types of statistical analyses involving test scores have produced interesting results, as well. Recent research along these lines is being used to refute the testing



movement. In a study by Cunningham and Sanzo (2002), the authors attempted to determine whether Virginia state exam scores were related to SES (socioeconomic status). “All high schools in Virginia were analyzed using subsidized lunch percentages and adjusted pass rates for the English, math, science, and history SOL [Virginia Standards of Learning] tests for 1998 and 1999” (p. 67). Their correlation regression analysis of the 245 schools provided evidence of a correlation between SES and test scores, with poor students’ scores falling below students’ scores with a higher SES. They concluded that “these results question policies that hold teachers and students in communities with lower SES to the same standards as those in communities with higher SES, and it challenges the practice of applying sanctions to students, educators, and schools on the basis of test results” (p. 67). In a similar study, Beck and Shoffstall (2005) analyzed Illinois junior high test data and found strong associations between SES and the Illinois Standards Achievement Test; they then extended beyond this correlation to include rural schools’ specific data in relation to SES and test performance.

Mahon’s (2006) study involving ELL students and high-stakes test data from Colorado investigated the question, “To what extent does English language proficiency predict English academic achievement, as measured by the CSAP [Colorado Student Assessment Program] in reading, writing, and mathematics?” (p. 481). She claims that there have been past comparisons of minority group achievement on high-stakes exams but little research on ELLs. Mahon used English language proficiency scores and CSAP achievement scores from 200 ELLs in the fourth and fifth grades. She used a variety of statistical analyses in SPSS (bivariate linear regression analyses, Bonferroni’s multiple

comparisons, a post-hoc power analysis, a one-way analysis of variance, etc.). Mahon's conclusion was limited and she decided that the results "cannot unequivocally establish the relationship between language proficiency scores and CSAP scores" (p. 492) though she did find particular relationships within certain analyses.

Carnoy and Loeb (2002) investigated the effects of strength of accountability programs in the 50 states on NAEP score gains. They found a correlation between high state accountability programs and NAEP eighth-grade math score gains among Black, White, and Hispanic students. Interestingly, they found "no evidence of a positive effect of accountability on student progression through high school" (p. 322). Lee (2006) conducted a similar study examining "50 states' activism in test-driven external accountability policies and their support for key school resources" (p. 43). Examination of NAEP fourth- and eighth-grade reading and math revealed little effects from accountability policy alone but possible significant effects related to availability of school resources. The author suggested further research into the area of inadequate resources and whether or not NCLB mandates are consistent with the lack of federal funding to make up for state resource inadequacies.

These examples of quantitative studies involving test scores are insightful if analyzing possible effects of multiple variables on test scores, and comparisons between scores on different testing instruments, but they cannot verify the value of what has been learned and tested, and they do not reveal the inverse effects of the tests on the curriculum and teaching. In recent years, scholars have begun focusing their attention on the topic of testing effects on curriculum, instruction, teachers, and students with

increasingly widespread interest. What follows is a summary of the past two decades of research on testing effects.

#### Testing Effects Research: 1990-1999

Though standardized testing has been in existence in varying formats and degrees for nearly the past 100 years, it took time for researchers to embrace the testing movement as something to be studied in depth. Zancanella (1992) cited a review of the literature on testing published between 1977 and 1987 in which the researchers found only 11 articles dealing with consequences of testing. Ten years later, Mehrens (1998) reviewed the literature on consequences of assessment; though a decade had passed since Zancanella found so little published research on testing effects, one of Mehrens' main conclusions, among others, was that the available research in 1998 was still inadequate.

In the 1990s, studies on testing effects increased considerably over the previous decade. These studies revealed important insights about the effects of standardized exams on curriculum and instruction and suggested that future similar studies would be worthwhile and informative. The following studies provide an overview of what has been found by researchers. Paris and Urdan (1994) investigated teachers' attitudes about testing by surveying 153 K-8 teachers. They found that teachers felt negatively about standardized tests, and that teachers admitted to manipulating their practice to improve students' test scores. Herman and Golan (1993) surveyed upper elementary teachers from nine states looking for "effects of standardized testing on schools and the teaching and learning processes within them" (p. 20). They found, among other things, that teachers

felt pressure to improve test scores and adjusted their teaching and curriculum due to this pressure. Smith (1991) reported a qualitative study in which her group's research included observation, interviewing, and document analysis over a 15-month period. The group found that testing led to reduced instruction time, narrowing of the curriculum, ignoring nontested subjects and significant changes to instructional practice in weeks leading up to standardized exams.

Sullivan (1991) effectively utilized both interviews and surveys with teachers and found that tests do affect classroom practice directly and indirectly. Bond and Cohn (1991) surveyed teachers and administrators in a study where most participants believed that any effects on the curriculum would be positive due to a refining process. Glasnapp, Poggio, and Miller (1991) surveyed school board members, teachers, and administrators and found that testing effects were weak except in lower-performing districts. Brown (1993) interviewed 42 teachers and principals and found distrust in state testing and the policy-makers and that testing may be oversimplifying curriculum. In a multi-year study in which data was gathered through interviews (with state-level policy makers, teachers, and administrators), observations, and document analysis, Grant (1996) found that tests influence curriculum development, cut into teaching time, and influence classroom assessment; in addition, some of his more recent research involving focus group interviews over a 2-year period with teachers suggested that while tests do affect instruction, this is dependent on a variety of factors (Grant, 2001). Like Grant, Zancanella (1992) found that teachers believed tests affect curriculum, but this was also dependent on other factors (like teachers' roles and personal views, and the amount of support they

receive). Zancanella's research centered on case studies of three teachers and included interviews with teachers and students, observations, and artifact collection.

After surveying 39 math, English, and reading teachers, Tittle, Kelly-Benjamin, and Sacks (1991) reported teachers feeling little benefit from large-scale testing because tests were not used effectively. Stake and Theobald (1990), using surveys and group and individual interviews of teachers, found both positive and negative influences of tests on curriculum, with teachers claiming less pressure on them than is placed on students by high-stakes tests. Wilson and Corbett (1991) combined "qualitative fieldwork" (focused on interviewing a wide variety of staff members at 12 school districts) with follow-up surveys and found that as stakes increased, the tests had an increasing effect on instructional programs. Finally, Firestone, Mayrowetz, and Fairman (1998) interviewed teachers in five school districts in Maine and Maryland and reported one state's test (Maryland) having significant influence on curriculum while the other state's teachers (Maine) reported very little influence by their state test, demonstrating that variation from state to state does exist.

#### Testing Effects Research: 2000-2008

Reviewing the literature on effects of high-stakes testing on curriculum and instruction up through 1999 answers, at least to some degree, the question, "Does high-stakes testing have an effect on curriculum and instruction?" The educators being given voice seem to be telling us, "Yes, tests affect the classroom in a variety of ways." Over the past eight years, there has been a substantial contribution to the research involving

teachers' and students' perceptions of the effects of high-stakes testing. Curren (2006) claimed that this increase in high-stakes testing interest came with the NCLB legislation in 2002 and the subsequent proliferation of high-stakes testing in the United States. However, even as recently as 2005, researchers are still claiming a dearth of research "about actual impacts of high-stakes testing" (Pringle & Martin, 2005, p. 349). Though the number of qualitative studies incorporating teacher and student voices in the national dialogue on testing effects over these past 8 years is relatively small (and arguably, inadequate for making general conclusions), the recent studies looking for teachers' and students' perceptions of testing effects claim important insights. The following articles are samples of the work being done in recent years.

Au (2007) analyzed 49 qualitative studies associated with how testing affects curriculum. He used a method called qualitative meta-synthesis, which allows a researcher to synthesize a group of qualitative studies "to gain a better understanding of the general nature of a given phenomenon" (p. 259). His primary finding was that high-stakes testing has "the predominant effect of narrowing curricular content to those subjects included in the tests...compelling the teachers to use more lecture-based, teacher-centered pedagogies" (p. 264). However, among his other findings, he noted that in some cases, researchers found that high-stakes testing caused an expanding of the curriculum and student-centered pedagogy. His conclusion was a question: "Are test-driven curriculum and teacher-centered instruction good or bad for teachers, students, schools, communities, and education in general?" (p. 264).

Crocco and Costigan (2007) interviewed dozens of teachers in New York City.

Their qualitative study “focused on eliciting narratives about [teachers’] work” (p. 518). Over a 5-year period, they conducted over 200 interviews with beginning teachers (all teachers interviewed had no more than 5 years of experience). In addition to one-on-one interviews, they held dozens of focus groups. After gathering a large body of data, they looked for emergent themes and wrote their conclusions. Among their several conclusions, they suggest that one of the “unintended consequences of the accountability movement in NYC’s public schools may be the narrowing of curriculum and pedagogy” (p. 529). Teachers revealed that testing produces mandated or narrowed curriculum and the perceived feeling that a teacher has lost control of the practice of teaching. For many of these teachers, this loss of curricular control was an impetus to leave the profession.

Valli and Buese (2007) spent 4 years studying the effects of high-stakes testing on teachers. Approximately 150 teachers from 25 schools were interviewed as individuals or in focus groups. The study also included administrators and educational specialists. The study focused on analyzing teachers’ roles and, in particular, tracked the “policy directive” of differentiated instruction to follow role changes in light of testing pressure. Their most dramatic finding was summarized: “Because teachers and principals felt so pressed to implement so many changes, they seemed unable to prioritize instructional improvement efforts according to the needs of their students or themselves” (p. 553). Though the study revealed adverse effects of high-stakes accountability, their conclusions also included the possibility that teacher role changing has positive potential if educators approach policy directives and changes carefully and thoughtfully.

Watanabe (2007) conducted ethnographic case studies and interviews in a study

that revealed a narrowing of the curriculum, among other findings, due to high-stakes accountability. Watanabe focused on middle school educators in North Carolina. Data was gathered over the 2001-2002 school year, and analysis included looking for emergent themes in a grounded theory approach. The study focused on language arts instruction. Among the emergent themes, the study found that testing reduced teaching time, affected student enjoyment of reading, reduced cooperative learning, negatively influenced authentic writing opportunities, pushed teachers to focus on products instead of process, separated writing from literature, and narrowed aspects of the curriculum. Watanabe found that testing displaced priorities at all schools.

Paris and Urdan (2000) reviewed the findings of several surveys conducted among teachers in three states. They found that teachers felt negatively impacted by high-stakes testing. The survey results suggested that high-stakes testing undermines teacher innovation and creativity, and that high-stakes testing has serious negative consequences for students. Based on their findings, Paris and Urdan did not call for the removal of high-stakes testing; rather, they outline policies for the better use of assessment in schools. They agree that teachers should be held to high standards, but they suggest there is a better way to use testing in this effort.

Massey (2006) conducted a case study of a first-year teacher and the influence of high-stakes testing on her literacy instruction. One of the study's guiding questions was, "Because Paula was hired to teach in a 'gateway' grade (meaning that students would not be promoted unless they passed the test), would testing affect her literacy instruction, and if so, how?" (p. 73). Research included interviews, transcripts of emails and phone calls,



lesson plans, and field notes. Massey concluded that this beginning teacher was negatively influenced by high-stakes testing, but that teacher educators, mentors, and school leaders can help new teachers adapt to high-stakes influences. This puts a positive spin on a phenomenon many scholars and educators see as a negative.

Abrams, Pedulla, and Madaus (2003) published an article that gave a brief overview of the literature on teachers' perceptions of testing programs. They also conducted a nationwide survey of teachers, gathering opinions on the influences of testing on teaching and learning. They found that teachers generally perceived a negative influence on their instruction but this was somewhat dependent on the "nature of the consequences or stakes attached to their state test results" (p. 22). In this survey, teachers were generally positive about state standards, but most teachers felt that the high-stakes state tests narrowed curricular focus and diminished the breadth of content.

Pringle and Martin (2005) researched the impending impact of high-stakes testing on elementary science teachers in Florida. The authors surveyed 38 teachers using closed and open-ended questions. They found that teachers had significant concerns about finding test preparation materials, shifts in curricular focus, the unknown of the tests, whether students would be successful or not, and the amount of time needed to prepare children for the tests. Many of the teachers' schools had already sent them to training for curriculum changes. The general anxiety and fear of the teachers was clearly evident in the survey. Oddly enough, many of these teachers' fears were centered in their own lack of familiarity with science and science standards and a comfortable reliance on textbooks. In a similar, but larger, study, Jones and Egley (2006) surveyed 708 teachers and 325

principals and assistant principals in Florida using open and forced-response questions. Generally, they found that teachers, more than administrators, had negative responses toward the state testing program. The most notable conclusion was that teachers believed that the state testing had no positive influence on student learning, but it had tremendous influence on teaching.

Sloan (2007) reviewed some of the ethnographic research which provides qualitative insights into the effects of high-stakes accountability on teachers and their classrooms, with particular emphasis on minority youth. This adds to a predominance of conclusions from statistics and research studies that minority youth are especially hit hard by high-stakes testing (Borg, Plumlee, & Stranahan, 2007; Brennan, Kim, Wenz-Gross, & Siperstein, 2001; Causey-Bush, 2005; Escamilla, Mahon, Riley-Bernal, & Rutledge, 2003; La Roche & Shriberg, 2004; Valenzuela, 2005; Wiley & Wright, 2004). Sloan (2007) cited Smith's (1991) claim that we need to get beyond teachers' verbal opinions about testing and get into their classrooms regularly to see the real effects. The ethnographic studies Sloan reviewed provided clearer, more vivid pictures of the effects of high-stakes testing. For example, the Pennington study he reviewed gave insight into reputations and feelings of teachers, students, administrators, and even the community. Classroom and school-level strategies, failures and successes, and changes in curriculum materials were described, elucidating the actual effects of the high-stakes testing on a school community through firsthand observation. These ethnographic studies provided powerful descriptions of the significant testing effects on teachers and minority students.

Lattimore (2005) interviewed six African American students to gain their

perceptions on preparing for high-stakes mathematics testing in Ohio. Lattimore's work was founded on a dual claim: "Those who understand the effects of testing are those who are closest to the test—the students. Ironically, students' voices are rarely heard in the testing debate" (p. 138). He focused "on African American students because testing seems to disproportionately disadvantage low income students of color" (p. 138). Each student was interviewed six times. The author found that these students were not so much afraid of or pressured by the test; he found a "surprising lack of bitterness" and students even claimed that "the test provided an impetus to try harder in school" (p. 143). The negativity students had was toward boredom in the classroom. Lattimore suggested this was due to mediocre teaching which may be influenced by the high-stakes tests and the need to prepare for them.

Lloyd (2007) studied a student teacher's kindergarten teaching experience in a high-stakes accountability climate in an urban, low-performing school. The student teacher was observed and interviewed over her 10-week student-teaching internship. Classroom artifacts were also collected. According to Lloyd, "the main analysis phase of the study focused on identifying factors that influenced Bridget's design of mathematics instruction" (p. 335). The study found that the student teacher was significantly influenced by the directives of the school, which were very test centered. Worksheets and structured lessons replaced socialization and developmental activities, and this teacher had to consciously plan beyond the school's workbooks to allow for more activity among the students. The student teacher repeatedly made conscious decisions about whether to follow a school-promoted, test-centered curriculum that was very skill oriented or to add

her own elements that she seemed to feel might be somewhat at odds with the school's style of teaching and curriculum.

Triplett and Barksdale (2005) studied 225 elementary school children's perceptions of high-stakes testing through drawings and written responses to questions. On the day after a high-stakes test, the children drew pictures and wrote about their testing experience. The authors categorized the drawings with the most common category for drawings being "emotions." Two of the most frequent emotions were "nervous" and "angry." Many of the drawings had "unhappy and angry facial expressions.... Smiles were nearly nonexistent" (p. 245). Interestingly, "teachers were nearly absent in students' drawings and writings (only 19 of 225)" (p. 248). Other categories for the drawings included, Easy, Content Areas, Student Metaphors, Fire, Power/Politics, Adult Language, Culture of Testing, and Not Enough Information. The authors were surprised at the overwhelming negativity and felt that the findings "convincingly support previous research reporting that elementary children experience high levels of nervousness, worry, and anxiety about high-stakes testing" (p. 255).

Van Hover, Hicks, and Irwin (2007) investigated the following question in a qualitative study involving seven beginning high school history teachers: "How do beginning teachers conceptualize historical thinking within a high-stakes testing context?" (p. 92). The study centered on interview data but also included observations, email communication, and classroom documents. The teachers were interviewed individually, and the interviews were semi-structured. Teacher opinions varied about the influence of high-stakes testing, but all were concerned about "SOL [Social Science

Standards of Learning end-of-level tests]—imposed time pressures—specifically, the rush to cover ‘everything’ before the SOL tests in May” (p. 96). Classroom observations verified the time pressures teachers brought up in the interviews, and the researchers concluded that “the SOL test appeared to exert a pervasive and tacit influence on how teachers decided to teach.... And this meant, in the teachers’ minds, that in order to do their jobs properly, they had to cover an enormous amount of content at a rapid pace” (p. 109).

Grant (2001) conducted case studies of two high school social studies teachers in New York to “explore the influence of state-level testing” (p. 399). The author observed every lesson in a unit taught by each teacher and conducted semi-structured interviews with each teacher before and after the units. Grant approached the study on the premise that “the research around teachers and tests [of which he claims there is not enough] shows no clear pattern of influence. Tests seem to matter, but how and to what extent is unclear” (p. 402). After interviewing and observing the two teachers, Grant concluded, “I see little direct, deep, and consistent influence of tests on these teachers’ classroom practices. The pervading sense that tests drive content, instruction, and the like seems alternately overstated, ill informed, or misplaced” (p. 421). He further suggests that “considerations of personal biography..., local organizational structures..., and the state and district policy climate may also figure into the content and instructional decisions teachers like Strait and Blair [the two he studied] make” (p. 421).

One of the more interesting and in depth studies of high-stakes testing is reported in the book, *High Stakes: Poverty, Testing, and Failure in American Schools* (Johnson &

Johnson, 2006). The authors, Dale and Johnson, stepped aside from their university teaching to spend the 2000-2001 school year as third- and fourth-grade teachers at a poverty-stricken school in Louisiana. The authors maintained a journal of their teaching experience, and the book combines these daily journals with the “voices” of the students, their fellow faculty members, politicians, and school policy leaders. Their choice of Louisiana for this qualitative study was important: “In 2000, Louisiana became the first state in the nation to require elementary (fourth grade) and middle school (eighth grade) students to pass a standardized test for promotion to the next grade” (p. 201). In summarizing their experiences as teacher-researchers in Louisiana, they note several important findings:

Three major themes are addressed throughout this book: the grinding effects of acute poverty on all aspects of life, the negative consequences of the continuing drive for accountability in the schools, and the unreasonable demands placed on teachers that stifle their creativity and enthusiasm and hasten their exodus from the profession. (p. xviii)

The authors’ perspective and findings are important to the dialogue on high-stakes testing because they were not just gathered within the school setting, but these themes were experienced first-hand. They claim high-stakes testing has a profound negative impact on teachers, the students (especially the poor), the curriculum, and the school in general.

Costigan (2002) interviewed six first-year third- to fifth-grade teachers in the New York city area. The teachers were interviewed at the end of their first semester of teaching and again in the middle of the second semester. He found that testing quickly became a primary concern among these teachers and had an impact in the following ways: a negative impact on students; a negative impact on classroom practice; a feeling

among the teachers of being unprepared for the volume of testing; a negative impact on the quality and nature of the instruction; and a feeling of powerlessness among the teachers. Among Costigan's final conclusions, he said that for these teachers, "a very real culture of testing has been created in the schools and districts in which they teach, and they feel they are unable successfully to negotiate between a testing curriculum and personal best practice" (p. 33). He suggested a deepening of the dialogue between these teachers in how to teach within a testing culture.

Escamilla, Chavez, and Vigil (2005) investigated high-stakes test scores in Colorado and teachers' perceptions of Spanish-speaking ELL (English language learners) and Latino students' achievement outcomes and results at these teachers' schools. The authors gathered descriptive data about the K-12 Latino and Spanish-speaking population in Colorado, led discussion groups of 35 teachers and summarized emergent patterns from these group discussions, and gathered data to investigate the question, "Is there a gap in achievement between Spanish-speaking ELL Latinos and other students at selected urban schools that are highly affected by linguistic diversity?" (p. 135). Among various findings, the authors noted these two important conclusions: First, the teachers perceived the Spanish speakers as responsible for "their low and unsatisfactory ratings on the state's school accountability reports" (p. 138); Second, the state test data shows Spanish-speaking Latino students in ELL programs who take the Spanish version of the state tests are meeting state standards and performing quite well. The authors challenge the teachers' perceptions that Spanish-speaking Latino students are underachieving and point out "that these same educators are unable to identify evidence that may counter this

perceived reality” (p. 142).

Considered as a body of work, these studies reveal a strong set of implications for further research into high-stakes testing effects. Some of the research suggests the effects are minimal, but generally, the studies reviewed here suggest high-stakes testing negatively affects various aspects of the public school system and the students this system serves. As the research on high-stakes testing has expanded over the past decade, an increasing number of studies have determined that philosophical claims about negative testing effects have at least some merit. Of these studies reviewed here, those that sought teachers’ voices, or at least the perspectives of practicing educators within the school system, seemed to have the strongest suggestions that high-stakes testing does affect our curriculum and instruction. At the very least, these studies reveal a culture of testing has established itself in our public schools; it is a culture that teachers are trying to understand and make sense of as they strive to educate our youth.

Reading these studies, it becomes apparent that a more concerted effort should be made to access the expertise of those on the frontlines of education—the teachers who carry out the policies, adjust the curriculum, prepare the students, and translate into practice the ideologies driving our schools. We need to hear more of what these practitioners have to say about what is happening in our classrooms regarding curriculum and instruction as a result of high-stakes testing (Cimbricz, 2002). These teachers are the point of focus as they carry a load of directives, policy, tests, accountability, and standards into their relationships with students. These teachers are the filters through which all public school policy and procedure regarding student achievement passes. Their



views of how testing policies affect curriculum and instruction are vital to developing a greater understanding of testing effects. Qualitative studies giving direct voice to teachers can make significant scientific contributions to a complex dialogue in need of those voices most intimately connected to the curriculum, instruction, and testing practices of schools.

### AIMS Test Research

Important to this study's topic is the fact that no studies have been conducted to determine teachers' perceptions of positive or negative effects of the high-stakes AIMS test on curriculum and instruction in Arizona. In fact, there is only a handful of published literature about AIMS testing. Ganesh (2002) obtained Arizona teachers' feelings about the AIMS test by having teachers draw how they felt about the test. While his findings are indeed unique and interesting, they do not tell us what influences testing is having on curriculum and instruction; and, it should be noted, his study was conducted before the AIMS became high stakes.

Wright (2005) analyzed the three policies guiding curriculum reform efforts in Arizona and the potential effects on ELL students. The three policies were *AZ Learns* (the program that includes AIMS), NCLB, and Proposition 203. The analysis was policy language based, meaning the author was looking for overlap, exclusion, and contradiction and how the cumulative effect of these three policies might affect ELL student populations.

Smith (2005), a prominent researcher and professor at Arizona State University,

included a brief political history of the AIMS test in Chapter 2 of her book, *Political Spectacle and the Fate of American Schools*. Smith suggested that state superintendent of education, Lisa Graham Keegan, who established the AIMS test as Arizona's high-stakes exam, was influential on a national scale: "[Keegan's] highly visible AIMS program provided a model for the national testing policies adopted by presidential administrations of both parties and fit perfectly with George W. Bush's educational platform" (p. 67). In this chapter, Smith includes an important question and assertion that adds to this dissertation study's relevance and importance: "Have tangible benefits of assessment policies to students, schools, and society been documented? Real effects are hard to find, and states themselves rarely even look for them" (p. 72). As researchers seek answers to questions and challenges like Smith's, seeking out teacher perspectives will add to our understanding of testing effects and the assessment policies that may be affecting our "students, schools, and society."

### Conclusion

Dewey's views on curriculum and the art of teaching establish a particular view of pedagogy that informs this study. Dewey's work provides a theoretical foundation for educators who are opposed to, or at least concerned about, the development of standardized schools and curriculum, particularly over the past half century with the proliferation of standardized testing. Reviewing the past four decades of literature on high-stakes testing effects and the history of standardized testing over the course of a century exposes the need for a broader and deeper body of research on the possible

effects of high-stakes testing. High-stakes testing has become a pervasive and contested part of our schools in the U.S., and the extant testing-effects studies provide a solid foundation of insights and possible profound effects upon which future studies might build for a clearer understanding of this important national issue. Scholars have examined the issue of standardized testing in schools through the use of philosophical analysis, quantitative studies, and qualitative studies. A deeper analysis, utilizing all three of these modes of inquiry and analysis, would serve U.S. education well.

Both qualitative and quantitative (especially surveys) methods have provided a sound body of research on which to build. Out of these studies, several important ideas arise. First, high-stakes testing does affect teacher's perceptions. Second, these perceptions suggest that high-stakes testing is affecting curriculum, instruction, and our students. Third, research on high-stakes testing, while improving in breadth and depth over the past three decades, is still in need of a broader body of work to improve our understanding of testing effects. In particular, more research that gives voice to teachers and more research that examines the effect of the AIMS test on curriculum and instruction is needed. This study contributes to a growing field of research on high-stakes testing by adding a new perspective to a body of foundational studies related to high-stakes testing effects.

## CHAPTER III

### METHODS

The methods of inquiry for this study focused on three principles: case study, focus groups, and educational criticism. I will begin with a review of Eisner's ideas about qualitative inquiry and educational criticism to establish the theoretical foundations for this study's methods of inquiry. Second, I will review the principles of case studies and focus groups. Finally, I will detail my methods for participant selection, data collection protocols, maintaining credibility and trustworthiness of the data, and acknowledge my limitations as a researcher.

#### Theoretical Foundations

The theoretical foundations for the research methodology used in this study rest with Eisner and his ideas on qualitative research and, in particular, his original thinking on educational criticism. Throughout his writings, Eisner repeatedly made the case that “we are inventing new ways to conduct research” (1994, 1998a, 1998b, p. 108). He claims that the top-down approach of scientifically analyzing teaching and using the data to adjust practice can have purpose, but he sees educational research not “as dependable prescriptions for action but as analogues to increase the quality of teachers’ deliberations” (1998b, p. 112). Eisner (1994) drew on Dewey in describing teaching and educational criticism as arts: “‘the function of criticism,’ wrote Dewey (1934) ‘is the reeducation of perception of works of art (p. 324).’ The critic’s task in this view is not primarily the issuance of a judgment but rather the difficult task of ‘lifting the veils that keep the eyes

from seeing (p. 324)' (Eisner, 1998b, p. 213). In his writings, Eisner asks every educator to gain a clearer picture of what is happening in our schools, to open our vision to the art of teaching and curriculum, and to develop an informed dialogue about what is really taking place in the classroom (1998b).

Educational criticism (for a thorough definition see Chapter 10 of Eisner's, *The Educational Imagination*) is "making public through language what one has seen, interpreted, and appraised;" and for Eisner, educational criticism "may be the only way in which the subtleties in teaching can be revealed" (2002, p. 57). "Criticism" wrote Eisner (1994), "is the art of disclosing the qualities of events or objects that connoisseurship perceives" (p. 219).

In Eisner's view (2002), criticism has four major aspects. First, the descriptive aspect is the attempt to communicate what one sees, but this is probably the most difficult part of criticism because "the critic's verbal magic must be most acute" (p. 227). The educational critic must describe a new vision in realistic but vivid terms. This skill of critical writing is one that Eisner believes can be developed just as skills in writing in the more academic mode can. However, Eisner's educational criticism has a more aesthetic quality to it and often resembles literature more than traditional research writing.

Second, the interpretive aspect is somewhat more dependent on the expertise of the critic because it involves seeing the context and finer qualities of a situation so that she can make interpretive conclusions. In describing an aspect of education, one must interpret to a considerable degree, especially in the field of qualitative inquiry. Eisner (1994) related interpretation to Geertz's (1973, as cited in Eisner, 1994) "thick

description” where the more details one observes, the greater one can deduce or interpret (p. 230).

Third, the evaluative aspect “most clearly distinguish[es] the work of the educational critic from that of the social scientist” (Eisner, 1994, p. 231). In this aspect, the researcher appraises the observed phenomenon. Eisner considered this evaluation “vital” and said, “to describe students’ work, or the processes of classroom life, without being able to determine if this work or these processes are miseducational, noneducational, or educational, is to describe a set of conditions without knowing if those conditions contribute to a state of educational health or illness” (Eisner, 1998a, pp. 99-100). Again, the expertise of the critic is important here because an evaluation has more power and meaning when the researcher has higher levels of perception and understands the context of the phenomenon.

Fourth, the thematics aspect of criticism can be summed up with a question: “What does it all add up to?” (Eisner, 1994, p. 233). Thematics is basically extrapolation and projection. What does this phenomenon tell us about the larger context or society in general? Thematics is a generalization made based on the observation of a smaller situation. “Every classroom, school, teacher, student, book, or building displays not only itself, but features it has in common with other classrooms” (Eisner, 1998a, p. 103). If a critic sees and appreciates qualities of a situation beyond the normal observer, then the generalizations of this critic are more likely to be trusted.

Eisner believed researchers need to be invested in the educational world they are investigating. After listing a number of general educational researcher questions, he

writes, “Questions like these are not to be answered by examining new methods of instruction or by scrutinizing test scores. They require an intimacy with what goes on in schools. Schooling needs to be ‘known’ ...by direct, intimate, contact” (Eisner, 1998a, p. 11). The goal of educational criticism is to develop new understanding and to help us to see what we may not have seen before.

Driven by Eisner’s ideals for qualitative inquiry, this study sought to inquire into the effects of a high-stakes exam, not through analysis of test scores, or through a distant survey of teachers, or through simple document analysis; this study’s purpose was to conduct research that builds upon teachers voices in school settings, discussing real, immediate testing issues and effects, and to get into the classrooms of these teachers where their expressed perspectives could inform observations and document analyses. Additionally, I was able to compare and contrast their expressed perspectives to what I was able to observe, and what I discovered through the study of documents. In this way, the perceptions of the teachers was respected and incorporated into the study but my status as a critic was also incorporated into the study as I developed an analytical and aesthetic understanding of their views and my own perspective throughout the conceptualization of the study, the review of the literature, the development of the methods, and the data gathering and interpretation process. I sought to incorporate the four aspects of educational criticism—description, interpretation, evaluation, and thematics—as much as possible throughout the study.

Eisner repeatedly addresses Dewey’s focus on experience as education in his works. I believe Eisner was driven by this notion in his development of educational

criticism and the need for qualitative educational research that critiques the artistry of teachers and learners. Eisner wrote, “Experience, for Dewey, is the means through which educational processes work, hence understanding education requires appraisal of the kind of experience individuals have” (1998a, p. 99). This case study into the effects of a high-stakes exam on curriculum and instruction, where the experiences of teachers is sought and observed, is motivated by such a philosophy and intended to be a valuable contribution to the high-stakes testing dialogue.

### Case Study

This research was a case study centered on focus group interviews but also included classroom observations and document analysis. This was a multisite study involving teachers of math and English at three Arizona public high schools where the high-stakes AIMS test was a required part of the curriculum. The boundaries of the study were limited to these three sites, the math and English teachers at these three sites, and their classrooms. Two Native American schools were chosen because, despite the existence of tens of thousands of Native American students on multiple reservations in Arizona, educational research in Arizona historically ignores these native populations and their schools.

The purpose of this case study was to obtain a group of Arizona teachers’ perceptions of the effects of the high-stakes AIMS exam on curriculum and instruction. The intent was to understand, from the perspective of the teachers, how these influences are actualized in the classroom setting. Classroom observations and document analyses,



informed by particular perceptions teachers had about AIMS effects, provided additional data that could be used for triangulation purposes.

Questions this study focused on are as follows.

1. What are teachers' perceptions of the effects of the high-stakes AIMS test on the high school English curriculum?
2. What are teachers' perceptions of the effects of the high-stakes AIMS test on the high school math curriculum?
3. What are teachers' perceptions of the effects of the high-stakes AIMS test on teaching practices?
4. How do teachers view the relationship between these effects on or changes in curriculum and instruction, and their power as teachers, and the power of their school administrators and state office officials?
5. What subtle, relatively unseen effects that might only be seen by a teacher or student in the actual classroom do teachers perceive?
6. What can classroom observations and document analyses add to the researcher's understanding of teachers' perceptions of AIMS effects?

It should be noted that my interviews were not restricted to the above questions. These questions were asked to stimulate discussion, but I often asked follow-up questions dependent upon, or influenced by, participant responses and the flow of the interviews. For example, during one particular conversation at High School C, the teachers were discussing the idea of shifting curriculum due to test score influence and preparing students for the AIMS. I asked the following question: "If the test wasn't there and you

did have a batch of students who were struggling, would you shift things around more if the test wasn't there?" I remember wondering at the time if the teachers were blaming the AIMS for shifting their curriculum around just because we were talking about the AIMS. I felt that the question would give them an opportunity to evaluate their remarks and clarify their points. I did not do this often, but there were times when I added a question to stimulate the conversation. There were also times teachers asked their own questions, which I share in the interview excerpts in the following chapter.

An important outcome of this case study is its contribution to the relatively small, but growing field of qualitative inquiry into testing effects on the classroom, particularly the AIMS test. In qualitative research, one reads about the need for "thick description" to support the drawing out of valid inferences (Creswell, 1998). The testing effects dialogue generally is in need of a thicker description and the expansion of the body of qualitative research on testing effects will contribute to our understanding. Authors, both those decrying high-stakes testing and those applauding high-stakes testing, need more data. Teachers, and the policy-makers who influence teaching, need a clearer vision of testing effects to, as Eisner and Dewey would claim, improve our perception and educational understanding. Investigating teacher perceptions of a genuine high-stakes state exam like the AIMS can provide a solid contribution to the testing dialogue. Finally, the verity of the oft-voiced concern that "the test will become the curriculum and that instruction will be narrow and focused on facts" (Domenech, 2000, p. 17) can be more fully understood when we include the perceptions of those professionals who develop and implement the curriculum—teachers.

I chose to conduct a qualitative case study because the method is particularly well suited to the goals of the study. After reviewing the literature on testing effects, one quickly arrives at the question: How can we gain further insight into the effects of high-stakes tests on curriculum and instruction? While surveys can provide an overview of respondents' views, the relatively short list of researcher-generated possible responses are limited (Bangura, 1992) and they provide a limited range of information. Because a case study involves multiple sources of qualitative data it is a viable and effective method of research for accomplishing this goal.

Creswell (1998) described the basic principles of a case study, which would include the following: (a) a description of the problem or case and the situation in which it is grounded—the context which is “bounded by time and place” (p. 61); (b) multiple sources for data collection to develop clarity with the issue or problem; (c) possible inclusion of personal experience within the context to inform and interpret the data gathered and with which to draw conclusions, which might include “an analysis of themes or issues and an interpretation or assertions about the case by the researcher” (p. 63).

The study by Wilson and Corbett (1991) was an example of the benefits of using a case study to obtain teachers' perceptions of testing effects by using teacher interviews. They began by interviewing staff in 12 school districts, generating a survey based on these interviews, and then conducting follow-up interviews after the survey. The studies involving open-ended interviews are able to provide a broader range of insights. Giving teachers a prompt and letting them discuss the issue or question opens the door to both

expected and unexpected responses that come directly from the participants.

Grant's (2000) study, "Teachers and Tests: Exploring Teacher's Perceptions of Changes in the New York State Testing Program," provided an excellent example of the data that can be collected using case study interviews, or specifically, focus group interviews. Interviews were conducted over a 2-year period. The first year they interviewed two focus groups made up of elementary and high school teachers representing "a cross-section of urban, suburban, and rural school districts in western New York state, a breadth of teaching experience (2-25 years), and a range of school subjects (language arts, mathematics, science, and social studies)" (p. 3). The second year they repeated the study with the same teacher participant demographics, with some new participants in each of the two groups. The author used a variety of methods to code and interpret the data (constant comparative method, looking for emergent themes). The author summarized his findings as follows.

Looking across the interviews, I saw patterns which help explain the teachers' responses in a social context and the nature of their learning in an array of social settings. The three preliminary patterns I synthesized from the data and report on in this paper relate to the nature and substance of the tests, the professional development opportunities available to teachers, and the rationales for and the consequences of the state exams. (p. 4)

Grant's (2002) summary of the implications of this study states that teachers' opinions and views must be considered in state level planning and reform. While this study did not specifically look for teachers' perceptions of testing effects, the structure of the article and the guiding principle that teachers' perceptions of state intrusions on the curriculum need to be ascertained, was important to developing a rationale for my study. This study adopted, with some variation, Grant's methods of inquiry (focus groups), and a version of

his methods of data analysis (identifying emergent themes and looking for correlation between interviewed groups).

### Focus Group Interviews

Focus groups have gained a solid footing in the field of qualitative inquiry. Most notably, focus group interviews are used extensively in the world of business, especially marketing, and in recent years have gained wide acceptance in the field of health, especially nursing (Goebert & Rosenthal, 2002; Halcomb, Gholizadeh, DiGiacomo, Phillips, & Davidson, 2007). In addition, focus groups are an accepted form of research in the fields of psychology (Reis & Wheeler, 1991), communications (Morrison, 1998), cultural anthropology (Bryant, 2007), and law (Ball, 2001). With the wide acceptance of focus groups in these and other fields, researchers have begun to set forth practical application of the method in the social sciences (Kleiber, 2004; Seidman, 2006).

An important contribution to the field of focus group research is Morgan's (1997) book-length paper, *Focus Group's as Qualitative Research*. He wrote a defense of focus group interviewing as a viable method of qualitative research. In addition, he summarizes the uses of focus groups; how to design a focus group interview and study; how to conduct a proper focus group as a researcher; how to analyze interview data; and he presents an overview of possible affecting factors a researcher might encounter while conducting focus group research.

Equally important, and more detailed in its treatment of certain aspects of focus groups, is Greenbaum's (1993) book, *The Handbook for Focus Group Research*. He

included important chapters on being an effective focus group moderator. The researcher who moderates a focus group needs to avoid particular mistakes to eliminate bias and improve validity. Greenbaum's approach is more business or marketing oriented, but his ideas are easily applied to an educational setting.

Jayanthi and Nelson's (2002) book, *Savvy Decision Making: An Administrator's Guide to Using Focus Groups in Schools*, provided an important summary of the effective use of focus groups in schools. Their book outlined effective strategies for focus groups in school settings and provided a sound argument that viable data can be acquired. This book, together with ideas in Kleiber's (2004) work on focus groups in the social sciences, provided much of the procedural foundation for the focus group methods used in my study.

### Participant Selection

This study centered on several focus group inquiries into teachers' perceived effects of the high-stakes AIMS test on high school English and math curriculum and instruction. Three high schools in three Arizona school districts were chosen as sites for the focus group interviews. The schools each had student populations between 500 and 800 students. Two of the districts are located on the Navajo Reservation with 99% of the student population Navajo. The third district is located in a rural, predominantly white community. I chose to include two Native American schools to emphasize a point and to make a particular contribution: though there are thousands of Native American students in Arizona, they are almost always left out of the studies that contribute to state

educational policy analysis. School districts' and principals' permission to conduct the interviews was secured, and IRB approval was obtained for the study through Utah State University and the Navajo Nation.

The focus groups were drawn from high school English and math teachers in three schools. The focus groups consisted of between 4-8 people (the literature allows for a variety of focus group sizes, usually between 4-12 people). A total of three focus group interviews lasting approximately 90 minutes each were conducted at High School A, High School B, and High School C. A balance of English and math teachers and male and female teachers at each school site was sought. In addition, a balance of Navajo and non-Navajo teachers was sought in the two Navajo high schools. These balances happened naturally among the groups and no interventions were needed on my part to ask for more volunteers to create these balances. The only imbalance was at High School A where we only had one male participant in the focus group.

The first priority was to select volunteers who were experienced teachers, preferably 5 years or more in an Arizona high school English or math classroom. The intent was to draw participants who taught prior to the AIMS becoming high-stakes and during the 3 years of its high-stakes status. Only English and math teachers were involved to set some limitations on the scope of the study and because English and math are the only current high-stakes sections on the AIMS test. I served as the only moderator for the focus group interviews.

## Data Collection Protocol

The interviews were largely open ended. I used the previously noted list of six questions (with a few possible variations) to prompt discussion. I only stepped in as moderator if I observed the conversation drifting from the central goal of identifying teachers' perceptions of the effects of the AIMS test on English and math curriculum and instruction. As moderator, I avoided providing answers; the purpose of this study was to clear the way for teachers to provide their direct perceptions to important questions about the effects of high-stakes exams.

The interviews were recorded using a digital recording device; these interviews were then transcribed by an assistant. Data were analyzed using a variety of methods suggested by Miles and Huberman (1994) in their book, *Qualitative Data Analysis*. The data were coded based on emergent themes or meaning. As I read the three school site transcripts, I created a separate list of themes or ideas that seemed to emerge from the text. Following the first reading, I took the list of themes and grouped the list to reduce the categories. These groups resulted in six emergent themes. I then read the interviews a second time and color-coded participant responses based on the six themes. At the completion of this process, I had six themes with interview excerpts clustered under these six headings. These clusters of themes are presented as a narrative summarizing the ideas within the participant responses. In addition, I share important contrasts or contradictions between the three sites "that sharpens understanding" (Miles & Huberman, p. 245). Miles and Huberman suggested that "conclusion drawing" was evolving throughout the data collection and analysis process (p. 11). This study followed such an emergent path



focused on two analysis principles: finding and sharing commonalities between sites and participants, and identifying points of contrast between sites and participants.

In addition to the focus groups interviews, I conducted at least five classroom observations at each school site following the interviews to “see” the claims interview participants made about AIMS effects, and to open up further dialogue with each teacher in his or her classroom setting. It is understood that these observations do not provide a thorough substantiation of interview claims; rather, the observations were intended to provide another layer of description, or analysis, of effects the AIMS might be having on curriculum and instruction. I also asked interview participants to provide documents that substantiate claims made in the focus group interviews. Again, this documentary evidence is not thorough proof of claims, but another layer of evidence to improve the scope of the inquiry. The types of documents that were analyzed included all or part of the following: classroom assessments, lesson plans, curriculum maps, class assignments, teaching materials, samples from teachers’ professional libraries, and teaching guides. I took detailed notes during the observations and compared these notes and the collected documents with the interview transcripts looking for correlation or contradictions.

### Credibility and Trustworthiness of Data

Credibility and trustworthiness of the data was addressed through the triangulation of data by conducting interviews in three separate districts. As expected, emergent themes at one site corresponded to, and overlapped, emergent themes at other sites or what Morgan (1997) referred to as “group-to-group validation” (p. 63). In

addition, in that the study intended to investigate teachers' perceptions, I acted as a moderator who merely provided a few questions as impetus for discussion and no answers, so that the direct perceptions of teachers remained intact and will, hopefully, be recognized as credible representations of their inferences about testing effects. Finally, producing three forms of data (interview transcripts, observation notes, and teacher-provided documents) improved my understanding and credibility and trustworthiness of claims and inferences. Validity in qualitative research is achieved through a methodically and artfully constructed narrative of defensible, systematic, and complex processes rather than through the careful control of variables that is the hallmark of quantitative studies (Wolcott, 1994).

I used several methods to take my personal biases into account and to establish internal validity. First, I used a bracketing interview, answering questions from another researcher and articulating my thoughts and opinions as I described my own biases and positions regarding high-stakes testing. This allowed me to see my biases and deal with them appropriately as the study moved forward. Second, I employed a journaling technique as I moved through the process of the study, describing my thoughts going in and throughout the study. Third, I provided the opportunity for member checks of the interview data. Transcripts were sent to study participants, and they were allowed to review their responses and suggest possible revisions (Creswell, 1998). Finally, I invited a colleague to analyze my data and findings to verify my conclusions.

I also maintained a field journal. This included research activities, important developments, contacts, and personal introspection as my understanding and opinions

developed with the project. This provided a record of my data collection procedures and gave me a clear picture of my own development throughout the study. I have come to understand the importance of recognizing and tracing one's own biases and beliefs in a qualitative inquiry.

### Limitations and the Role of the Researcher

In this study, I acknowledge the following possible limitations. First, I relied on volunteers to come forward; volunteers are often those with strong opinions and this might run counter to the idea that valid inferences are best made on data from a randomly selected group. Considering this, I set several participation prerequisites (experience, ethnicity, and gender) to counter this possible limitation. Therefore, I hoped that volunteers had not only strong opinions on the subject, but that their views had partly been shaped by their gender, ethnicity, and particular experiences.

Second, as researcher, I created and asked the questions for my interviews which may have led the discussions down particular paths the teachers might not have ventured if asking and answering their own questions. I attempted to counter this with the focus group method of collecting data where participants are given greater freedom to develop the conversation and build on or react to the group's responses. As a data collection instrument, I attempted to remain unobtrusive and to guide the conversation only as necessary in keeping us on the central topic of teachers' perceptions of AIMS effects.

Third, I identified the emergent themes and assigned them importance. I may have passed over important ideas teachers shared despite my best efforts to systematically

analyze the data for emergent topics and themes. Although my perspective may have been a limiting factor, my experience as a teacher, an administrator, and an educational scholar strengthened my ability to develop a nuanced interpretation of the data.

Fourth, I have been a secondary education teacher for 8 years, and a school administrator for over 2 years. At least five of these years were spent in two high schools in two different states with high-stakes exams affecting graduation. I spent 4 years as an English teacher at one of these high schools preparing students for the exam. At another school, I spent 1 year as a Dean of Instruction responsible for evaluating the teachers preparing students for the AIMS exam, creating the tutoring programs for students who had failed the AIMS, and administering the test itself. As an educator with strong feelings from my own perceived effects of standardized testing, I may have introduced elements of this bias into aspects of my study—the question development, the interviews, and the data analysis and conclusions; though, I will say that going into this study I wanted to eliminate this fourth potential bias, and I have worked carefully to present teachers' true perceptions whether they be for or against standardized testing. In many ways, my experience as a teacher added to my ability to access the participants and the language they use and to take the perspective of the teachers in this study more easily.

Finally, I recognize that my views about standardized testing are open to dispute because of the lack of research available to support either side of this issue as well as the complexity of the issues surrounding this topic; I am pleased that some of my own preconceived notions of testing effects have been refuted through the completion of this study. I have intentionally searched for perceptions or data that counters my expectations

or that fell outside the obvious themes that emerged (discrepant data). I believe the data that follows will show that many of the teachers held similar opinions to mine, yet there are also important opinions that run contrary to my preconceived notions of AIMS effects and high-stakes testing effects in general.

## CHAPTER IV

### RESULTS

This chapter begins with a summary of the participant school demographics to provide a context for the data. Second, for clarification, I briefly review important educational terms used throughout the interviews by participants. Third, I set forth the six themes that emerged from the interviews. These will be presented in a narrative form, focusing on the teachers' voices with direct quotes but also including my commentary and opinions. This section of the chapter will include some analysis of the data that will be presented more extensively in Chapter V.

#### Participant School Demographics

At the outset of the presentation of data collection results, a summary of the school demographics is included to inform the reading of the interview data and help with seeing the comments of the teachers in the context of their school situations. Tables 1-3 show a summary of the demographics for each of the three schools. I have not included a summary of the interview participant demographics (gender, race, years of experience) because the departments are small enough at each school that the participants could be easily identified, and I have promised to share this study with each district while also promising participants anonymity. In order to retain school anonymity, I have changed the numbers to fit in a range (for example, instead of saying a dropout rate was 8.7%, I report 5-10%). I have also left out the internet source for the school report cards that report AIMS scores because the website address would identify the school. The three

Table 1

*Participant Schools' AIMS Scores and AYP Status*

School	2008 AIMS scores (% passed)			2008 AYP determination
	Reading	Writing	Math	
A	40-50	35-45	45-50	No
B	40-50	40-50	50-60	Yes
C	70-80	70-80	70-80	Yes

Table 2

*Participant Schools' Demographic Information*

School	2008 % free and reduced	2007 4-year grad rate	2008 dropout rate
A	80-85	70-80	7-10
B	80-85	60-70	7-10
C	25-30	80-90	0

Table 3

*Participant Schools' 2008 Ethnicity Percentages*

School	% Hispanic	% Native American	% White
A		99	
B		99	
C	15-20	1-5	70-80

schools are identified throughout this study as High School A, High School B, and High School C. The two schools on the Navajo Reservation are High School A and High School B. All of the information is reflective of the 2008 school year (the year in which

the interviews were conducted) except for the 4-year graduation rates, which are reflective of the 2007 year, the most current provided by the ADE at this writing. All of the information in these charts is available through the ADE website (<http://www.ade.state.az.us/>).

### Clarification of Important Terms

It is important to define a number of terms before reporting the data and explain how the schools and teachers were identified. As already noted, there were three high schools participating in this study (High School A, High School B, and High School C). The teachers at each school were each assigned an identifying number with the letter representing their school following the number. Thus, Teacher 1 at High School A is identified as Teacher 1a; Teacher 1 at High School B is identified as Teacher 1b, and so forth.

It is important to explain some of the terminology used in the interviews. Teachers at High School A and High School B talked about the restructuring the English and math curriculum has undergone in recent years, mainly driven by the AIMS. Teams of teachers have met with district curriculum specialists to conduct a “vertical alignment” or “spiraling” of the curriculum. This may have included teams of teachers, from kindergarten through grade 12, breaking down standards by grade level and subject area (in these three schools much of the initial focus has been English and math because those are the subjects tested in AIMS). Additionally, this vertical alignment may include assigning skills and content or subject matter to each grade level. The idea behind this is,



if a child moves through this system, all of the important or required standards will be taught to each student; they will move through district-determined grade-level appropriate content and skills systematically; and they will be prepared for the state exams, especially the graduation-required AIMS in high school. Among the perceived positive effects for those who espouse this system of education, teachers can reduce the amount of time spent reviewing or teaching skills, concepts, and content that should have already taken place at previous grade levels.

In addition to vertical alignment or spiraling based on grade levels, sometimes schools will assign standards, skills, or content to specific quarters or blocks within a school year. For this study, the breaking down of standards, skills, and content into quarters, semesters, trimesters, or blocks within a specific grade level will be called “curriculum blocking” or a “blocked curriculum.” Occasionally, the term “curriculum mapping” was used in the interviews, especially at High School C. This typically refers to a curriculum overview for a department and its courses. This might include a breakdown of POs (performance objectives) or standards by quarter or semester, with an overview of content, skills, assessments, and materials, as well. Curriculum mapping can be prescriptive (the curriculum is mapped prior to delivery) or retrospective (the curriculum is mapped after delivery to reveal what actually happened in the classroom). In all three schools, any “mapping” seemed to be prescriptive. Curriculum mapping may have different layers of details for each teacher, school, or district, but it generally parallels, or coincides with, curriculum blocking and vertical alignment efforts.

POs are the same as standards. In Arizona, standards, or POs, come from the

ADE, although, some high schools will develop their own, or synthesize multiple standards or POs as “power standards” or “power POs.” Power standards or power POs may also refer to single standards (rather than a synthesis of several standards) which may be represented in a high percentage on the AIMS.

Some form of this restructuring of the curriculum was taking place in each of the three high schools; however, the teachers at High School C referred to their changes as “curriculum mapping” (they didn’t use the terms blocking or vertical alignment), and they attributed their changes to a new district administration that brought the mapping concepts to their district. High School C teachers did connect the standards and mapping to the AIMS in the interviews, but they were not as persistent and direct in tying one to the other as was the case in High School A and High School B.

Before reviewing the themes emerging from the interview data, a brief summary of the AIMS test and how it is delivered to students will add to our understanding. The AIMS is first taken in the sophomore year in high school in the spring. If a student passes one or more of the three sections (reading, writing, and math), then he or she does not have to take that section again unless they are trying to achieve a higher score. A student may achieve one of four score classifications (from highest to lowest): exceeds, meets, approaches, or falls far below. A student passing a test with a Meets might try testing again to achieve an Exceeds. Students achieving exceeds on all three test sections may receive a scholarship to state colleges or universities. A student must achieve a Meets on all three test sections. Students may take the test again during the fall or spring of their junior and senior years for a total of five attempts. In the past, with certain stipulations,

the state has provided additional opportunities for some students in the summer or after their fifth attempt and completion of high school credit requirements.

### Emergent Themes in the Data

The focus group interviews were analyzed looking for emergent themes in the transcripts of teachers discussing their perceptions of AIMS effects. I identified six themes emerging from the data: creativity in the classroom; vertical alignment and structured curriculum; standardized curriculum versus nonstandard students; reviewing or cramming for the test; narrowing, expanding, and altering the curriculum; and cultural challenges and the AIMS. My intent here is to allow the teachers' voices to rise out of the interviews and provide qualitative data to inform the testing effects dialogue. The teachers in these groups had strong opinions about the possibility of AIMS effects. Some felt that effects were positive and some felt that they were negative. Some felt that there was a mix of the two. It was also interesting to see teachers coming to realizations and even questioning their own original feelings or ideas as they expressed themselves on the subject for seemingly the first time in such a formal setting with a defined purpose and outcome. What follows is a theme-by-theme presentation of these teachers' ideas mixed with my commentary about the interviews and my school visits.

#### *Creativity in the Classroom*

One of the themes that came out of the interviews was the idea that the AIMS was affecting creativity in the classroom—in lessons, teaching methods, and curriculum. Teachers said that they wanted to teach real life skills through creative, performance-

based activities and assessments, but they felt pressured by the AIMS to focus on other skills and knowledge. For some, the fun and creative side of teaching was slipping away.

At High School B, Teacher 3b introduced the idea to the focus group that there were other ways to teach and assess that were not predicated on the AIMS.

This type of test [AIMS] is not an indicator of kids' abilities. Some kids it is, but you can't have one test that's gonna have a baseline for all kids. You know even like when I was in high school we had to do a senior project. And that was a mixture of math and reading and writing, and we had to show you know similar to standards. We had to be able to show that we could put all of these skills together into one project and present it. It had a speech involved; we had to do community service hours. Those are the things that kids remember. Those are the things that tell us as educators, okay, this kid gets it and this kid has reading skills. This kid has writing skills; this kid can do basic math skills; not a standardized test.

Following this comment, the following exchange took place.

Teacher 5b: It's like they turn around, and they make it the kids' problem. And you know you don't get to go to these activities because you didn't pass...and so forth. And I think as a teacher, that's how my classroom changed because before, I was more creative based. I did performance-type assignments. They acted out skits and things like that, and it's like now I don't even have that in my classroom.  
Teacher 4b: But that stuff is real learning.

Teacher 5b: Yeah, exactly.

Teacher 4b: That's where they really learn; that's the stuff that sticks with them.

Teacher 3b: And kids don't go to their grandchildren and say, "I remember this great multiple choice question she made for me. I did so good on it. I remember question seven. It was awesome." I mean that's just not real life. They don't remember that stuff, but in my classroom, I have them practicing that kind of thing and that test taking skill, so that they can do better. You have to tailor to that.

Teacher 5b: It's really hard now to get creative.

Toward the end of this discussion, Teacher 4b referred to her senior English class and how different it was from her other classes because she could retain her creativity. "[In

this class] I have seniors, and I have a very few select that haven't passed the AIMS test, and I'm not gonna sacrifice creativity and authentic learning to focus on that. At that point they need to come in for tutoring." At High School A, Teacher 5a made a similar statement. She said that the school was working toward getting "the majority of tenth graders passing [AIMS]. Which means then an eleventh and twelfth grader are open for the fun, more difficult, more challenging stuff.... We're just not there yet, but we could be."

While the teachers at High School B were seemingly unified on the stance that the AIMS had reduced creativity, activity-based learning, and performance-based learning and assessment in their school, the High School A teachers had an interesting discussion on this issue that somewhat contrasted the discussion at High School B. The High School A teachers seemed to lean toward the idea that the AIMS did not inhibit their creativity, rather it forced them to seek creative ways to teach the standards. Early in the interview at High School A, Teacher 2a said,

As a teacher you have to become more creative in a way like how am I going to get these kids involved with hands-on activities, but at the same time, how are we gonna meet those standards? It's that creativity of the teacher. It's sort of extra work for us, but it's just something we have to face and hopefully meet.

Later in the discussion, Teacher 2a repeated the idea she had earlier expressed that creativity is improved: "In the past, I did a lot of lecture-based teaching and this time around I have to really let the kids do their own learning, have them interact with each other. So my lesson plans, they definitely changed."

Whenever the interview turned toward creativity in the classroom, the High School A teachers seemed to be almost defensive that creativity had not been lost, but

there were a few comments that suggested the AIMS was affecting their creativity:

Teacher 3a: We actually, I think are teaching more now because of the AIMS, not less. But you're right, with English, that was where all the creative units were when I was in school, you know. That was the fun projects, and you know it really does impact the time that you have for those kinds of things.

Teacher 2a: That's where the challenging part comes, when she said creativity. That's where we have to be creative—how can we implement fun activities? Geometry is supposed to be fun, so I have to make it fun; and we do a lot of colorful stuff—that's just to keep them going. But the thing is...there's a limited time I guess for that. I cannot do that every day. I cannot do that all the time. I have to scatter [it] out through the semester.

Teacher 3a: It's the short term creative activities you know. None of us can really, in theory, afford to take weeks to do this, you know design a project or whatever. We don't have much time to work with, but to have neat activities that take a day or two that you throw in throughout the year; it keeps the students interested and gives us a break... I think it's just that we do the mini versions of the creative stuff, instead of like these massive projects that used to happen. And I don't know that there's a better or worse.

Responding to the idea that teachers now insert brief, creative days into their semesters, Teacher 5a said this is alright with the school's administration; for some reason she felt the need to validate the practice: "I'm not experiencing the type of pressure that maybe other people are, where if it seems like if somebody walks by your classroom and it sounds like it might not be boring, that they're like, "What are you doing in there? Are you not working on AIMS?' We don't have that here."

As the teachers at High School A discussed creativity in the classroom, the English teachers began discussing creative writing and how, since AIMS, they have replaced it with the more formal 5-paragraph essay that caters to the writing portion of the AIMS.

Teacher 5a: You know, the creative writing thing, I think, that might be one of those things that frequently teachers love so much that they're blinded to how

students feel about it. And I feel like what's fun is to get students to be creative, to do creative writing, to tell stories within their five paragraph essays, whatever expository. And then it's almost like sneaking it in, sneaking it past those kids who are afraid of it.

Teacher 3a responded to this by pointing out that the school had found a way to keep creative writing for the students, although, they had made the decision to remove it from the core set of English classes required for graduation.

That is how our school has addressed it is that we are now creating an elective for creative writing, which gives you the kids who actually want to do it without making everybody jump through that hoop; so it's not something that a student has to sacrifice in terms of the AIMS. It's something they can do in addition to their courses that are Arizona standards.

Toward the end of this discussion on creativity at High School A, the following two comments were made by two English teachers. The first teacher was arguing that formal essay writing and the Six Traits (a rubric used to grade formal writing and the basis for evaluating the Writing on the AIMS) meets the creative writing needs of her English students:

Teacher 5a: I just feel awful being this crazy booster for like the AIMS, but the Six Traits that we have, I mean, Voice is in there. And creativity is part of each one of those, and just because it's not a poem or you know traditional short story, doesn't mean that we're not doing creative writing; they are.

Conversely, at High School B, Teacher 8b felt that the AIMS had adversely affected her rubrics:

And now with the push to get kids onto [the next grade in the] system and onto the curriculum that the state has put out with its standards and POs, and then how that's gonna influence AIMS. Your rubric, well, my rubrics, end up reflecting that instead of the more creative things I would like to look at.

As an interesting side note, I heard the idea of a "state curriculum" repeatedly in these interviews. To be fair, the state of Arizona has not created a state curriculum. The state

provides grade level standards and curriculum guides, but the state does not require districts to vertically align or block their curriculum. It was interesting, however, that teachers had seemingly subconsciously come to equate their specific school curriculum, aligned, mapped, or blocked, with an imaginary Arizona state-mandated curriculum. Such was the focus on the AIMS at these three schools.

The final comment in this thread of discussion was made by a veteran teacher who feels that the AIMS has pulled teachers away from those aspects of the English curriculum teachers were passionate about and back to the skills that AIMS dictates high school students should have.

Teacher 1a: Well, to me sometimes a teacher's passions get so far away from what kids need to learn anyway, that it is more that teacher's issue with what they have to decrease rather than what's in the best interest for kids' general learning. And I've seen that come and go over the years in my own education as well as here; and here in particular with the AIMS. I think so many teachers got away with doing whatever they wanted for so many years that if they felt that pressure to decrease their personal passion, it was probably for the good of the kids to be honest.

The teachers at High School C were somewhat like the teachers at High School A. They were a little hesitant to discuss the idea of creativity and fun in teaching and the curriculum slipping away due to changes brought on by the AIMS. But in the second half of the interview, they began to open up a little. Teacher 4c was the first to address the issue.

I think the premise of this was a good idea: to make the teachers accountable for their teaching, and a lot of them had been slacking, and then make students accountable for their learning. If those are the two basic ideas, I think it kind of does that, in a way. But, I don't know that the pressure that it puts on the kids, and the remediation, I don't think it was thought through at all when they passed the bill. It has put a burden on teachers and actually, in a little way, is taking away some of the fun of teaching.



There was some murmuring of agreement after that thought, and Teacher 4c continued.

Well, I like to have fun, and I know it's taken away some of my fun. You know what I heard from a couple of teachers that retired years ago when this all came out? We're getting out of here before we have to do this. You know, I think there might have been some teachers that don't want to do this. That's not the way they want to teach. Do you understand?

This opened the door for more open opinions among the group.

Teacher 5c: I guess if you want to talk fun, I'd rather give my students an assignment and evaluate their writing and see where they are at in my class. Maybe this is me being in control and playing God, but how well do I feel they are meeting the standards. And I guess that actually isn't good because it doesn't matter how well I feel they're meeting the standards but how well the AIMS graders feel they're meeting the standards. But going through the data, I guess is interesting. But fun? Teaching should be fun.

Teacher 5c continued, as she seemed to be enjoying opening up about the issue.

I feel a constant need to focus on problem areas, and I guess that's good, but I don't know. I have a project that for me is kind of fun. Students have to do research outside of class, and then they have to present in class. I don't want to give that up. I don't want to give up, *Romeo and Juliet*, but if our scores, I don't know, I may have to give up, *Romeo and Juliet*, and concentrate more on assigning another essay for them to write and rewrite.

One of the math teachers then offered his opinion about teaching, especially, teaching the struggling students in the early math classes in high school.

Teacher 2c: In pre-algebra, I'd like to do more real world projects and things—in a perfect world, where I didn't have to worry about an AIMS test with certain standards. I'd like to do things because a lot of kids that get in pre-algebra hate math; that's why they're in pre-algebra. They don't do it very well because they don't like it—a catch 22. I'd like to make math more interesting to them, but they take too long. It'd be more fun to them, but it would take too long, and I wouldn't be able to get to things I need to so that I can get [them to] algebra I.

Teacher 4c came back with another English example.

I've found I cut things like dramatic presentations, and where they make costumes and really get into it because it seems like I don't have time for that because they need to be reading and understanding. They need to be writing. That is what I'm

talking about—fun stuff. That’s what keeps kids interested.

Hearing this, Teacher 5c then made an important clarification. She said, “Presentation is a standard, but it’s not a tested standard.” Teacher 4c agreed: “Of course! Like public speaking and presentations, and just being confident, and being a person in front of people, that’s not tested at all. That’s what I’m talking about. Why shouldn’t those types of things count for something? But they’re not anymore.”

Traditional speech standards (like presentation, listening, speaking, oral communication) have largely been moved into the core English curriculum standards. Speech classes by themselves are often electives within the English department. In Arizona one-half credit of the four credits of English required for graduation must include, but are not limited to, the principles of speech and debate. However, even though core English teachers are asked to teach speech standards, many are cutting these things out of their curriculum because they are not part of the AIMS exam. And, as I heard several times from different teachers in these interviews, the English teachers are working double-time to prepare their kids for two out of the three AIMS exams (Reading and Writing), so special projects or units are being trimmed.

While the teachers at High School A seemed to generally agree that creativity and fun in the classroom could still be addressed by creatively preparing kids for the AIMS, the teachers at High School B and High School C were a little more open about their frustrations with the loss of creativity within their profession. Teachers are preparing students for a state exam, and more schools are adopting curriculum programs that guide this preparation. As Teacher 3c said it: “You still do the other things, maybe not as much

as you want, as much variety as you would have brought in, your personality type things you know. That percent that used to be you get to kind of do what you like—you don't have that nearly as much. You have to stick with what's going down the pipe.”

*Vertical Alignment and Structured Curriculum*

At High School A, the school has been undergoing a thorough restructuring of the math and English curriculum with vertical alignment teams, identification of state standards or POs, and a focusing of the curriculum on skills and concepts represented on the AIMS. Teachers have been closely analyzing the test scores and the curriculum for several years. Several of the teachers at High School A spoke of these efforts as a positive effect of the AIMS on math and English curriculum and instruction.

I began the interview at High School A with the question, “What are your perceptions of the effects the AIMS may be having, positive and negative, on your teaching and the curriculum?” The first response came from Teacher 1a.

I'll start with the curriculum. [AIMS] really was the impetus for vertical aligning the curriculum and bringing teachers from K-12 together in one setting to discuss what was being taught—the gaps that were there and the strengths of the curriculum. So that was the first positive I would say.

Teacher 2a agreed.

To add to that, on a...positive note, the curriculum, the fact that when the high stakes test was involved, it allowed us to really focus on the curriculum itself. And we knew that it was the curriculum itself, it had to be clearly defined. And it allowed us to focus on that rather than doing anything else. I mean we had a job to do and that responsibility, accountability, really inspired, I guess some of us teachers, to then put in more effort, especially me.

At this point, Teacher 1a qualified what was happening—this happened repeatedly in the

interviews at all three high schools. As the teachers discussed what was happening, many seemed to really be thinking about these effects and considering them deeply, possibly for the first time, and various teachers would question or reconsider what was actually being said. For example, following the comment above, there was a thirty second gap, and then Teacher 1a said, “I think there’s an assumption, too, that we strongly believe that the, uh, the standards are viable and, um, credible, um, in order to be able to do that.” Teacher 1a pointed out that the above comments were positive, if indeed the school believed the standards and the AIMS were viable and credible and these subsequent effects were then credible. Comments on this topic from High School B will be examined later in this analysis and we will see that those teachers felt some concerns about the defensibility of vertical alignment.

Teacher 3a brought the discussion back to considering the positive effect of the AIMS focusing a school on curriculum restructuring and vertical alignment.

There’s a much better flow between courses and that’s part of the vertical alignment, but in addition to that, it has driven us to look at what courses we’re offering and at what levels we’re preparing our students. And we’ve eliminated a lot of lower level courses that the upper classes were getting bogged down in that they were getting to the AIMS test and not prepared for it; so now we’re getting students into higher level courses at a younger age which is preparing them more. And we’re finding that the students are rising to the challenge. So we were under-challenging our students before this, and as a result of the pressure of the AIMS, I think we’re now more accurately placing our students in courses that push their limits.

At this point, one of the teachers began to speak of the structured, standards-based, vertically aligned curriculum that moved kids towards the AIMS as both a positive and a negative. This teacher seemed unable to decide if she liked or disliked these AIMS effects.

Teachers 4a: I just say the main thing is just accountability, especially from a teacher's perspective. It makes us accountable for what we're teaching in the classroom. Although, I feel like there's not that flexibility. You know, I just remember as a kid, I watched my teachers and I liked what they taught us. I liked the flexibility they had and now it just seems like we're so standard driven by this curriculum. You know we have to go by it just because we have to meet state standards. That's the only thing that I see as a drawback is just that you don't have that flexibility as a teacher. Then again, you can get more creative with what's in front of you. So, as an English teacher...I can work around it and be creative with what's given to me.

One teacher felt that the structured math curriculum, with vertically aligned courses, was problematic at times.

Teacher 3a: And it's hard in math because it's so compartmentalized... When our students are in geometry they tend to do well on the geometry part of the AIMS, but then we see that the scores are dropping in other areas that aren't covered in the course. The same thing with algebra; when they're in the algebra courses, their algebra portion of the AIMS tends to be stronger. But then other things tend to take the hit, which requires taking time from the course content [which has been defined in advance as part of a multi-year, vertically aligned, scripted program]...we all know if you're not using something on a consistent basis, you lose it and it doesn't mean the kids never knew it, but if they haven't used it for a year, they forget it. So, we do have to take time in algebra class to review everything, and that's not necessarily a bad thing, but it certainly alters the structure of algebra. It alters the structure of geometry because the content isn't pure.

Another math teacher at High School A saw these vertically aligned, standards-based courses as a positive because of the preparatory effect for upper-level classes.

Teacher 3a: Because it [AIMS] drives the curriculum with specific standards, especially in math, and because we're redoing when students take the courses that are preparing them for AIMS, the freshmen and sophomores are very focused on the Arizona standards [that] have prepared them to pass the AIMS. But that means that we have more students when they get to the higher level courses have those basic foundations that in the higher level courses beyond the AIMS, the teachers actually can go farther and do more because the kids have the foundation they need. Whereas, in the past, kids would get those higher level courses and because the lower level courses weren't as driven in the skills that the students needed to have, teachers were limited in the post-AIMS classes. I think they can go so much further now.

Much later in the interview, this same teacher had one of those self-reflective moments where, as she expressed ideas and evaluated the direction of the curriculum, she began to question those directions. As a researcher, it was interesting to see these moments among the participants where they began seeing things, seemingly, from a new perspective.

The negative is even if the AIMS drives them to be focused in their education, once they pass, they have a lot harder time staying focused on their education. If they passed their AIMS, they're done. When they're upperclassman...they already jumped through that hoop and because we do forget learning for learning's sake and focus on learning to pass the AIMS, they think, what's the point of the rest of my education?

At High School A, the English teachers seemed committed to the vertical alignment of the curriculum and the paring down this had brought to their curriculum. Out of the three high schools, High School A was the most openly invested in this style of education.

Teacher 5a: I do feel like, oh my God, there's so much, we've just gotta keep going. And I just throw lists at kids like, learn it. Look at this please! But I know as we revise our curriculum, I think we could fix that problem. It doesn't have to be that way by getting tighter vertical alignment because as it is now, our curriculum document is...maybe over ambitious. But we do everything every year, [but instead] we could review everything before AIMS, you know, as needed. I think we would have plenty of time if we were all acting in accord with each other like ninth, tenth, eleventh, twelfth...seventh and eighth at the same time. We wouldn't feel like, oh, I didn't cover everything that could be on the AIMS.

Following this comment, several of the High School A teachers expressed frustration with the “shut-it-down” attitude of students after passing AIMS. They said that the school was working on developing post-AIMS honors and AP courses that will, as Teacher 1a said, “Take the place of AIMS. AIMS is done, but this is my next scaffold.”

The final comment on the vertical alignment theme at High School A came from

Teacher 3a, and seemed to sum up the attitude of the teachers in this group.

[AIMS] is making education. It comes back to the vertical alignment. It's making our education a cohesive unit. Their education in a given class shouldn't be primarily based on the teachers likes, dislikes, and personality. It should be based on the content that's appropriate for that grade level, for that course. And I think for so many of our kids, there's just this sense of disconnectedness as they moved from the year to the other because their courses used to be so much about the teacher and not about the content or the kids in their grade level and their level of maturity.

As a researcher, I could see that the AIMS test was having a powerful effect on High School A and its redefinition of how and what to teach. The entire math and English departments at High School A were restructuring curriculum, courses, and the methods of content delivery. Most of the teachers at High School A had bought into these methods of education; however, there seemed to creep into the conversation some concern as they discussed, out loud, these ideas that had come to change their school. It was a powerful conversation because it caused these educators to consider whether or not these effects of AIMS are good or bad, or a blend of both.

The teachers at High School B said that the math departments in the district had, in recent years, completely outlined each grade level and math subject with POs broken down by quarter. The district had developed benchmark tests for each quarter, as well. The English departments in the district had begun doing the same thing, but not all of the high school English classes had been structured around POs at the time of this interview. The math classes in the district had also been vertically aligned from early elementary through grade 12; the English classes were nearing completion of this project. The teachers even provided me with copies of the printed and bound math curriculum POs broken down by grade, class, and quarters. The teachers at High School B had strong

feelings about this AIMS-driven effort to create a structured curriculum program.

Teacher 2b: In math, everything is driven by the AIMS. Curriculum is created through that. We have the POs. It's separated into quarters that we have to cover as much as we can in those POs. Aside from that we have the vertical teams. I guess English has that, too. Everything is driven by AIMS. It's just teaching to the test. The creative part is lost because we have to teach to the standards to a point that they have to learn it quickly because they didn't learn it in the lower years. So, it's just catching up, catching up, catching up.

This frustration with vertical alignment was voiced repeatedly at High School B. The teachers felt that reservation schools struggle to retain teachers, often have long-term subs, often have disruptions to scheduling (at least this high school did), struggled with ELL problems, faced deep poverty issues, and all of this led to many students who were far behind at each grade level and in every class. The teachers at High School B felt that vertically aligned classes and rigid curriculum guides with POs divided by quarter might be having a negative effect on kids because teachers had to push quickly and move through pre-determined curriculum outlines without considering the individuals who were sitting before them. As Teacher 3b put it, "What happens is you have a kid that can't add, and you're supposed to be teaching them multiplication and division." The following teachers expressed similar thoughts.

Teacher 1b: The department of education also tells us that we need to teach this PO. And you know your kids well, and you know your kids need the skills, too, before you can teach them that PO, but it doesn't give you any option to do that. Even though you know they need it, you can't do it. You don't have a say.

Teacher 2b: No matter how good the POs are—we've laid it out, we have samples to execute them all—but if it's not executed on a lower year...I'm not putting the blame, but right now I know we're doing it because the pressure is on. We're executing to the tee—like PO this week, PO this week, PO this week. But, as I said, no matter how good it is, the execution down there, it's not as good. Then, by the time they come up to the high school, we're lagging behind.



Teacher 5b: I worked at another school where I developed the entire curriculum based on the AIMS scores. We'd look at the blueprint and see what percentages of the POs [were on the AIMS]. Yes, the Power POs; we covered those more in the curriculum and we hit those more throughout the year...versus the other POs [whose] percentages are not as high in the AIMS test. So, you just rush through. You disperse here and there and you don't concentrate on them. And the thing on teaching to the test, it's sad.

One of the teachers who had worked on the curriculum teams at High School B noted that they had determined their POs in the same way Teacher 5b had done at another school.

Teacher 3b: In terms of vertical team, what they do is they have the Power POs, but they also look at the rate the percentages of certain standards come up on the test. So, maybe a reading comprehension question comes up 40% of the time versus something else only coming up 7% of the time. So, they would put that standard in more often because it comes up more on the test.

Teacher 6b felt that the unifying and standardizing effects of the vertical alignment and curriculum blocking was a positive rather than a negative:

One thing I do like about the AIMS is that we're starting to get teachers on the same page. For example, a couple of years ago, when you taught Algebra I, this guy could teach it any way he wanted; he could grade it any way he wanted; he could teach whatever he wanted. He could start at Chapter 1 and get to Chapter 2. This teacher might start at Chapter 1 and get to Chapter 10, be a tough teacher, give homework every night. You know, it was very disjointed; it was very free. Whatever a teacher wanted to do, they could do. Whatever grades they wanted to give, they could give. There was no continuity between Algebra I with one teacher and Algebra I with another teacher.

Oddly enough, the interviews never moved into a discussion of how teachers could retain the freedom to teach in a manner that was "free," with focus on individual students at the different levels of ability, yet somehow still retain at least elements of the vertical alignment and curriculum blocking. The conflict was evident throughout the interviews; as Teacher 6b said, "I do like the idea of standards; you know I'm not crazy about the AIMS test for graduation, but as far as getting all teachers on the same page....it's not so

much teacher variations anymore.” As I reviewed the transcripts, I was disappointed that the idea never came up. I was disappointed that the teachers never proffered ideas on how to resolve the dilemma. However, I realized that the purpose of this research project was to discover these gaps in the discussion, these unresolved problems and, hopefully, introduce these topics or ideas to the conversation about testing effects.

Another effect of the AIMS that seemed to surface in these discussions was the idea that the AIMS, as a looming standardized, high-stakes exam, had created a desire amongst teachers to diagnose and categorize. They had been trained, if not formally, then informally by the AIMS, to categorize learning by grade level and even by quarter—they looked for a clean and scripted curriculum to tell them exactly what a group of students had been exposed to, taught, and experienced, so that they could then provide the next package of skills and information in the progression toward the AIMS.

Teacher 5b: At least you have some kind of previous background knowledge on the kids we have. When they come into your classroom, you know which POs are covered, and you know you can build on top of that. There’s, you know, the scaffolding and spiraling and what not. I think that’s what we lack in the English department [remember High School B had not yet fully blocked their high school English curriculum, quarter by quarter, year by year] because we don’t have that curriculum. And this year, I know for me anyways, it was really hard to assess the kids. You know, did Student A read, *Huckleberry Finn*, versus Student B read, *Of Mice and Men*? And it was just really, like how you were saying, disjointed, and I was really trying to figure out which kids had the background knowledge, and which kids didn’t you know.

This idea that the vertically aligned, blocked curriculum would provide a clear map toward the AIMS was pervasive among the teachers interviewed and classrooms observed at High School A. Teachers were constantly meeting in professional development—“our professional development schedule next year is pretty much driven

by the needs [identified by AIMS]”—and teachers were focused on reworking the curriculum to meet deficiencies represented in AIMS score breakdowns [the AIMS provides group feedback as well on various test sections or standards so teachers and administrators can see group weaknesses and strengths]. This data-driven decision-making was perceived to be a positive change-agent at High School A.

One of the other things we've seen though is not only the data in specific concepts, but like from the [AIMS] spring reading/writing scores. When we graphed the students in each category—Falls Far Below, Approaches, Meets, Exceeds—from the last year to this year, the overall passing wasn't changing dramatically, but the number of students who are approaching changed drastically. And just to realize that we are doing a better job of getting our kids over that line, well...we are seeing progress. And the progress isn't always on who passes the AIMS, but it's on if you've less Falls Far Below and you have more moving into the Approaches category, you know we're getting in the right direction. But we wouldn't know that unless we took the time to graph those results and ask ourselves what that means.

The teachers at High School C referred to their “curriculum mapping” repeatedly, and their need to hold to the standards at each grade level to prepare their students for the AIMS, but their conversation never settled in on this topic like the other two high schools. There were some direct comments on the subject, such as Teacher 5c: “How many times have we had to do curriculum mapping? They started ASAP when I started here, and so we curriculum mapped and basically just aligned everything to the standards. We keep doing it, and it's been done.” Moreover, Teacher 1c talked about how she actually tracks the standards in her classroom.

And we hit all our standards, whether we hit them hard enough.... I mean I know I keep track of all the standards I hit, and how many times I hit them; so that I am sure to hit all the standards at some point in the semester. But that's the way I was taught. You teach to the standards, you know.

Responding to this comment, Teacher 4c said, “And we sit down as a department, and we

make sure that we cover everything. And maybe we don't need to do it every single year, but we know we're covering everything." The teachers at High School C talked about meeting frequently to analyze and discuss AIMS data, to discuss standards and mapping, and to organize AIMS tutoring. For example, Teacher 5c was quite upset at one point and said.

I mean, God! I'm sorry you don't know this. We're a good school! We're a good school! We receive data. Our principal breaks data out for us. We meet in the summer. We look at our data. We look at eighth grade data coming in, even more so this year with our low tenth grade scores this year. Is there a connection? We don't know, but that's all we have to go on. And so we look at this year's ninth graders' eighth grade scores. Incoming eighth graders, we will look at their AIMS scores. That's what we do, and that's what we've been doing for several years to see if we need to alter things in our classroom.

However, in the interview, they did not dwell on a discussion of their feelings toward any school-wide curriculum program as did the other two high schools.

The vertically aligned, blocked curriculum, informed by data from the AIMS and guided by ADE standards, was a central topic of discussion in the two reservation high schools I visited. Though their opinions about this curriculum program were mixed, the predominant concern at High Schools A and B was that this system might leave some kids behind and ignore the individual student. The next section explores this theme.

### *Standardized Curriculum Versus Nonstandard Students*

Arising out of the vertical alignment, curriculum blocking discussions was an important topic about standardized curriculum ignoring the idea that individual students are not standardized. This idea was particularly important to the teachers at the two reservation schools where teacher turnover, student turnover, long-term substitutes, and

special cultural influences disrupted the attempts to create a vertically aligned flow to the curriculum. The cultural influences will be discussed as a separate section in itself, but the following interview excerpts are examples of the perceptions teachers shared.

Teacher 1a: The plus [of the AIMS] is having taught here many, many years and in different reservation schools, the test does equalize the playing field for what we expect from teachers and for students. But the negative is kids aren't always equalized by the time they get to you, so you're dealing with all those discrepancies in their backgrounds; so it does place an extra burden on you emotionally.

One teacher blamed the AIMS for taking away the idea of teaching the whole student.

Teacher 7b: I think AIMS is structured so you're not teaching the student, you're teaching something that is required by government—okay, here's this, this is what they need to learn. But you're not looking at the whole student. Like I said, we have a lot of students who have other issues [the conversation referred to the cultural influences and disruptions in the vertically aligned, blocked-curriculum system], and we're not looking at that, and the school is not staffed to where we can teach the whole student. I think with the AIMS...we're just tunnel vision. We're just gonna do this, and that's how much we're gonna do. And we're not looking at the student, we're just looking at them as low test-takers. Okay, here's the test, you take this, there you go. And it's come to that, and I for one, I don't like teaching that way. I want the kids to enjoy coming to school.

For Teacher 6b, a vertically aligned, blocked curriculum with the AIMS test awaiting her students, forced her to focus on the group achievement levels of her students in deciding not only methods of teaching but even who to teach.

Well, sometimes there are those hands-on activities that really develop understanding you just really don't have a lot of time to do. And I think that's a big problem; we're so overwhelmed with so many POs, you know that's a big problem. We've got so much to teach. It's kinda like throwing it out there—if they get it, great. The ones that are close, we'll help you go a little bit further. And the ones that fall far behind [the lowest AIMS category students are clustered in is, Falls Far Below], well, see ya next year because we're going on to the next PO right now. We don't have time to slow down and develop those conceptual ideas in their brain, what's really happening; which is really what they need. I guess we could break the class down into six different little classes and try and teach them all different things, but that's again, very labor intensive for the teacher.

Later in the conversation, this same teacher said:

I think I focus on those kids that are what you would call, Approaching [another official AIMS category for students just below Meets, or passing]. Those are the ones I focus on. And the exceptional kids, they've got it; and you try and throw them a bone here and there—a little more advanced for them while you're working with the other stuff. And then the FFB [Falls Far Below] kids, hmm, you know it's like, whew, you got so much work to do.

One of the disquieting occurrences in the interviews and in the subsequent conversations with teachers at the school sites during their classroom observations was the application of AIMS labels to groups of students and individual students. Students were repeatedly referred to as Exceeds, Meets, Approaches, and Falls Far Below. The slang had even developed into acronyms—"the FBB kids."

In speaking of the structured system of education the AIMS had influenced, Teacher 1a expressed the mixed feelings so many teachers had about the system.

I see the AIMS as kind of a Catch 22 in the long term. Like we need it to improve [this feeling of needing a test score to show improvement was pervasive at High School A] and to equalize the playing field, as I mentioned before, but by the same token, there will always be that group of kids who for whatever disability...or just timing—it isn't their time to know and understand yet—that they will be forever demoralized and penalized because of this. Yet, if we give it up then we'll decrease our gains of where we've grown and what we've been able to offer. And I'm speaking specifically on the reservation....the gains that I see we're making in this very short period of time between last year and this year even are huge. By that same token, there's still always gonna be that group of kids that I feel really sorry for, you know.

At High School C, the teachers spoke of students who wear labels or who get caught in the standardized system of preparing for AIMS. Teacher 5c said:

I don't know how you measure accountability in education because you've got people working with people. And yes there's definite objectives in math; you have to learn certain things in English, and you can measure that, but you're also dealing with people. And maybe it's just this first year, while we're first dealing with the fact that these two people won't graduate, but we're also putting a big

impact on their future.

Teacher 5c brought up an interesting dilemma facing their school. The teachers said that this was their first year where students were not graduating due to the AIMS. They did not linger on this, though, and quickly moved on with the discussion.

One of the topics related to this theme was brought up repeatedly at High School C. The teachers questioned whether or not their curriculum program was having any affect on the AIMS scores. Teacher 3c complained that the rise and fall of scores had more to do with the unique make-up of students at each grade level.

The thing that's not taken into account, though, is year to year you look at the groups of kids who come in. One year you have a fantastic group, and then all of a sudden, you get this batch of, you're trying, yeah. And it's not so much they can't, but the attitude of, I don't care. I'm not doing anything. Although you're bringing them in, you can't force it into the brain if they're not willing to meet you in there.

Unfortunately, some might just cast this comment aside as "excuse making," but this is a reality many teachers feel they are facing, and it is somewhat uncomfortable to those teaching in a standardized curriculum program with a high-stakes test looming at the end.

Therefore, the teachers at these schools did not necessarily dwell on this topic for a prolonged discussion. However, their comments were frequent enough, and the discussions at High School A and High School B were long enough that I felt this theme should receive some treatment here. The most important worry I heard was the idea that some teachers were frustrated by their blocked, vertically aligned curriculum that assumed the students were arriving each quarter at the same stage of learning with similar backgrounds of knowledge and understanding (nonstandard students in a standardized curriculum program). Within the core classes, this proved problematic for some teachers

who wanted to address individuals on their levels of understanding but felt compelled by the system to push on.

My visits to the schools for observations and further discussion with the teachers in their own classrooms revealed that High School A and High School C had developed strong tutoring and remedial programs for their students. High School C had an “early out” each Friday, but those students who needed tutoring remained for another period of school. High School A had developed classes around accelerated math and Galileo programs that helped those who were struggling with basic math concepts. High School B had some tutoring in place, and they did try to assign students who had failed AIMS to “AIMS Review” English and math classes that, while they were still core classes, were primarily made up of students who had failed the AIMS. High School B, though, was apparently struggling with the details of these efforts, and they were not as fully developed as the other two schools. The evidence was there, though, that all three schools were at least attempting to deal with the non-standard students in standardized classes through alternative, remedial programs supplemental to the core classes.

### *Reviewing or Cramming for the Test*

One of the catch phrases in the testing literature and testing discussions is “teaching to the test.” In this section I’ll share the teachers’ opinions about how the AIMS has caused teachers to devote time and curriculum to specific preparation for the AIMS itself, an idea already introduced in previous sections, but one which I will develop more fully here. Some of these teachers believed this was a positive, and some saw this as a negative. I’ll share teachers’ views on reviewing or “cramming” for the test



specifically, and also share some teachers' views on the idea of "teaching to the test" where the general curriculum or teaching methods were influenced.

Teacher 1c: I'm the sophomore teacher, and so that's the year that they take [the AIMS]. But the one thing I have to say about AIMS is at least it's a standards-based test. And it's not some let's pick some questions out of the hat. I mean, some of them seem that way. But for the most part, if you cover your standards, you're covering your test. And so you know you always hear the term, teaching for the test, and blah, blah, blah. I mean, I guess I do because I do a lot of persuasive writing. I do a lot of the stuff that has been on the test, but if you cover your standards... I mean it affects my classroom, yes, greatly, because it's the make or break year for my kids. Now, do I think it should be? Well....

I followed this comment by asking if teachers felt comfortable with this potential influence AIMS might be having. Teacher 1c responded:

I mean I'm very conscientious about my standards, and I'm sure to hit my standards. And so, in that respect, yes I am covering what I need to cover. Are there different ways that I think I could teach those same standards? But I feel like I need to word things the way that the AIMS might be worded so that way, the kids are used to the jargon that will be on the test. I mean, there's better things, better ways to teach things than what I'm doing, but I feel like I have to prepare them for the stuff they will see on the AIMS.

Teacher 2c pointed out that while the AIMS affects particular points of his curriculum, it does not necessarily control it.

I think in my math class I don't necessarily teach to the test, but if there's something that I know is emphasized on the test, I point it out to the students when it is being covered. I don't think it's changed what I teach, but I do emphasize points that are important.

At High School B, Teacher 8b discussed the thorough influence of the AIMS on her writing curriculum and her creation of a unit on questioning verbs to prepare students for test questions.

And it goes back to AIMS as well...I took a two-week period with my juniors, and we did a questioning verb unit. We went over all sorts of questioning verbs, and they hated it by the time we were finished. But they could answer the

questions. And we kept reviewing that all year. And it's reflective in their writing as well because if they can't read the question, if they can't understand the passage, they're not gonna write in the format they're asked to write in. So, between the questioning verbs and then really hitting up five paragraph essays and how to structure those in formal outlines for AIMS, it helps.

Teacher 8b was focused on writing, and she talked about a journaling activity where the students were actually required to write about their reviewing for the AIMS.

Before the AIMS [came] up, I took a month, and we took ten minutes out of the beginning of every class and did a test strategy. And some of them the kids knew, and they were wondering, why do we go over these? But it was a good review for them to get thinking about how are you going to address that multiple choice question that you just cannot figure out. How are you gonna look for main ideas in a paragraph or in a short story? How are you gonna look at true-false questions? And I had my kids keep journals about how they did, and it was really amazing how many had never been told that this is how tests are formatted and this is why.

Writing seemed to have been strongly influenced by the AIMS in the English department at High School B. Teacher 5b shared how she prepares her students for the AIMS writing test.

I show a lot of papers that were written in the past by students that are already graded by the state department. I really didn't do that before. And just to show the kids the examples, and they then go about rating their paper themselves and do their editing in that same fashion. So, I guess using model papers from the state department is something that I have incorporated into my classroom in the last couple of years just because of the high stakes.

However, writing was not the only subject strongly influenced by AIMS at High School B. One math teacher, Teacher 2b, shared a typical day with his pre-algebra students. Most of his classes were made up of students who had failed the AIMS in their eighth grade year (the AIMS is administered in lower grades to give schools a gauge of where the kids stand as they prepare for and enter high school where they will take on the AIMS required for graduation).

I teach pre-algebra, and these are the kids who just went FFB or fell far below during their eighth grade AIMS. So, mostly I have freshmen and a few sophomores, juniors, and seniors [who have not passed the high school AIMS or need credits]. To address the gap, I just teach to the test because of the pressure of the AIMS. What I do when I introduce a PO is I get a sample AIMS material from the actual AIMS test that was given a year or two ago from the ADE. I show that, have them answer off the bat without anything, just that's the first thing. Some of them will be able to get it, some will be close, and some will just be way off. And then I teach the PO; bombard them with examples boom, boom, boom, boom, boom, boom, boom, boom, boom, boom, boom, boom, boom. And then, at the end of the day or after two or three days, I show it again. Some still get, some get it, or most of them get it. So, that's the style I do just to address the gap. So they get used to the kind of test they'll be exposed to, instead of being on the creative side. But as I said, this is pre-algebra, and I have to address the gap. Just get them used to seeing the kind of questions they will encounter.

I observed this teacher, and the classes I observed flowed the exact way he described them in the interview. Talking to the other teachers, this teacher was generally regarded as one of the finest teachers in the high school and probably the best at getting the “kids who just went FFB” to pass the AIMS or at least out of the Falls Far Below category and up to Approaches.

At High School C, an interesting discussion focused on test preparation in the curriculum and its actual effects. Teacher 5c said one of her main frustrations was that after preparing her ninth grade students for an entire year, she did not really know if the preparation had affected their scores or not.

As a freshmen English teacher, I try to provide the foundation for the AIMS test for their sophomore year. And what I find really frustrating is that - I teach in their freshmen year - they take the test their sophomore year and scores come back, and again it's subjective. I don't know what I did in my classroom that caused or didn't cause a student to be proficient in AIMS Writing or AIMS Reading. Our scores are down this year after being up for the last two years. I couldn't tell you what I did or what we did as a department or a faculty to cause those scores.

I thought this was one of the most interesting statements of the entire interview because it

seemed to give voice to a recurring frustration or at least a concern (and the level of concern varied from teacher to teacher) at each of the three schools. Each school was following a curriculum designed, to varying degrees, to move students toward success on the AIMS. While some liked the structure of this curriculum, others were concerned about whether buying into this vertically aligned, blocked, or mapped curriculum was necessary, or what influences it was actually having on AIMS scores.

I asked the teachers at High School C if there was a shifting in the curriculum, year-to-year, influenced by the AIMS scores because of the rise and fall of scores that had been pointed out. Teacher 5c responded:

Well, in years past it might be influenced by my own boredom or professional interest, and I might take a summer school course on some aspect of literature and want to bring something like that into the class. Well, if I do that now, I've got to make sure it's tailored around AIMS, so that it will target the AIMS skills, or else don't bother bringing it in.

While not connecting the curriculum changes to test scores, Teacher 1c did say that her fall semester now ended in an intensive test-preparation review session, and the beginning of her winter/spring semester began with the same preparation session to get kids ready for the AIMS in February. She had moved the creative projects to the beginning and end of the school year.

Well, for me, it's just the fun is trying to get kids ready, especially second semester, in a month and a half. You know, so I am hitting them everyday with AIMS practice. This is the state standards; this is what it might look like, you know blah, blah, blah. And just having that timeline with that one group and then the other group. I do a fairy tale project, which is a lot of fun for me, and we do that earlier in the [fall] semester, which would be nicer at the end of the semester. But instead I am hitting persuasive writing and then turning around and doing the exact same thing with this new group [winter/spring semester] before they have to take the test. And so it does change my curriculum that way—what standards have been on the test; what standards have been pushed for you, like in Strand

Three in writing. That's what I hit real hard and you know the persuasive writing and the technical; and so instead of going from one thing to another and maybe being able to say, oh let's spend a couple more days on this and let's go into this a little bit deeper, if there's a class that really likes something, well, we've got that down, we better go on. Move and get this so that way you've got that basic knowledge because you know their sophomore year I have to hit everything.

This idea that teachers wanted to slow things down a bit, incorporate more of the creative and fun, and spend time where classes seemed to need to spend time was being pushed out of their teaching. Teachers at all three high schools felt like they were on a schedule, especially in the freshmen and sophomore classes, and they had to move along at a steady clip and cram and review in earnest as the AIMS appeared on the horizon.

Some of the teachers felt motivated by a fear of the AIMS to review and cram, sometimes just before the test and, for others, throughout the year:

Teacher 5a: I don't know if it's the test's fault, but I know just getting scared. It's scary. So sometimes you know you cram. You waste days in cramming. Like, okay, you get this prompt, what do you do? You know, and it's just drilling. And that's like, say, a few days of instruction, but add that up, right, and it doesn't seem right. Like, we should be ready to go; we should be fine, and that feels like a negative. The cramming that happens right before the test. I mean when the teacher personally gets afraid of what is about to happen.

Teacher 5a said she was specifically referring to the cramming right before the test (the AIMS is given in the fall and in the spring for juniors and seniors; high school students take it for the first time in the spring of their sophomore year). Teacher 4a responded to this and said the cramming and fear took place throughout the year, partly because she teaches ninth grade.

Just that idea of did I prepare for AIMS; did I prepare them enough? You know, it just freaks you out. You're constantly wondering, did I prepare my ninth graders for tenth grade? That's my big question every single day. And then like she said, you're cramming constantly. Oh man, they gotta know persuasion; they gotta know expository; they gotta know this and this research. I mean, you have to

constantly, I mean, I don't know. I just....

When I asked the teachers at High School C if they had changed their curriculum or teaching in any way to help kids deal with the fear or anxiety that they said AIMS brings to them and their students, there were several brief responses where teachers said their compassion for the kids had increased, or as Teacher 4c said, "I'm trying to be nicer" (this would be an interesting study in itself). The longer responses focused on the issue of teaching testing tips or testing practices that would help the kids succeed. Teacher 2c said:

I've done more multiple choice in my classes because of AIMS. They can get familiar with going backwards with an equation. Okay, I've got these four choices. One of them has to work, so I'm gonna go backwards, and I'm gonna plug in. And, oh, it equals—that must be this one. So those things have changed.

One response elicited laughter from everyone in the interview at High School C when I asked "How do you teach kids to deal with the anxiety of the AIMS test?" Teacher 4c immediately responded, "Padded cells."

One of the repeated frustrations in the interviews was connected to the idea of teachers trying to figure out what to teach and how to teach based on how the test would ask questions—teachers were frustrated trying to figure out what would actually get tested and how they could properly teach to it.

Teacher 1a: Well, and you were talking about all of us, our weak spot in literature is literary terms, and everybody's teaching literary terms, but I think the negative part of the test is figuring out, what do they want? How are they testing? It is so different from the way that all of us are teaching that we jump through that specific hoop to make sure our kids can test on literary terms. When our kids know how to use them; they know what they are, but apparently it's not in the same way they're being tested. So figuring out those little tricks, I think, is the negative.

Teacher 4a expressed a similar thought.

So, you know, it's just like for us as teachers, we're on a guessing game with AIMS every day. Trying to figure out, okay, maybe this will be on the test; maybe this will be on the test. No, this is not really important. So, with our curriculum, it's flexible; but with the AIMS, I don't know.

One of the outcomes of a vertically aligned, blocked curriculum is often what was referred to at High School A and High School B as "benchmark tests." Districts will work with companies to create these little AIMS-like tests, or they may have teacher teams and curriculum specialists create them within the district. The tests are designed to gauge student understanding of the standards blocked and taught that quarter as they prepare for the AIMS. Both High School A and High School B referred to preparing for these tests.

At High School B, there were as many as five benchmark tests a year, in addition to the AIMS and regular classroom tests. At High School A, Teacher 5a said:

I think the problem may be a problem with AIMS and sort of scrambling to meet our students' needs in terms of AIMS in the English classroom. They're tested all the time. And that's something we've talked about figuring out. And they feel it. And they're bored and you know they're tested on reading and writing all the time. Like AIMS-style tests, you know four times. Everybody has a benchmark, you know, every quarter. It gives them a perception of school as sort of boring.

At High School B, Teacher 3b had a similar thought.

We assess our kids—this year it was less, but last year we assessed our kids for six weeks out of the school year. With all of the benchmark testing for the AIMS and the actual AIMS test. And then that's not even including any of the assessments you do in your own classrooms. That's just district-wide assessments. I mean these kids are tested and tested and tested and tested and tested.

Talking to these teachers, they seemed frustrated that the AIMS was powerful enough in its effects on the curriculum and the reviewing or cramming that took place, but these benchmark tests extended this influence even further. The benchmarks enhanced the

power of AIMS to shape and guide their practice. They were also extremely frustrated with the claims that the benchmarks were a tool for teachers to better identify and teach to their students' strengths and weaknesses as the year went on. Rather, it became a tool for administration to monitor teachers' adherence to the curriculum map—the Quarter 1 benchmark will cover these standards or POs, so teachers need to cover those in their teaching. When asked if the benchmarks were used to gauge weaknesses and thus provide data for revised teaching the next quarter or semester, several teachers pointed out that in a vertically aligned, blocked curriculum, you did not have much time or freedom for such revisions. After the comment above from Teacher 3b about benchmark testing, the following exchange took place among the teachers.

Teacher 7b: But the thing about it is that we test, then it's like, okay, so then how are we gonna use that?

Teacher 1b: What happens next?

Teacher 6b: You don't have time to do the remedial work; you gotta move on to the next thing. So you identified that this kid has a problem.

Teacher 3b: And the next time they do the assessment, they have the same problem; and the next time you do the assessment, well, they have the same problem.

The last comments from the teachers at High School A on this topic seemed to be one of those reflective moments as they considered the opinions that had been expressed and their meaning. The teachers had just discussed the idea that they remind the kids about the AIMS all year to motivate them and to connect the learning to a larger purpose—the AIMS. There was an almost 30 second pause, and then two teachers spoke.

Teacher 1a: We focus on saying aloud to the kids [repeatedly connecting the learning to AIMS and reminding kids of the AIMS to constantly motivate them],



as opposed to learning. And that may be something that we as teachers have to work on, too, is I know I did it before AIMS was even high stakes, talking about the AIMS coming up, and you gotta learn this for the AIMS. Learning for the sake of learning, I think needs to be a conversation we have as teachers.

Teacher 5a: I think it should be a conversation we have with students, as well. Maybe particularly in English because I do have some students who talk about it a lot who will say, this is all we ever do at school.

Three of the teachers at High School B had a similar moment of questioning the current system of instruction as they discussed the topic of reviewing for, or teaching to, the test.

Teacher 5b had just shared her method of reviewing by making practice copies of AIMS-like tests, and she said, “I mean, you have to train them even from that level, just the endurance part of it, just having them sit still and concentrate that long.” Teacher 3b responded, “I think the key word there is train. We’re training kids instead of teaching them. So, this is not creating life-long learners; it’s not inspiring kids to enjoy learning, and teachers are not inspired to teach to the AIMS test.” Teacher 4b closed the discussion with this thought, “I was always taught that teaching to the test is not real teaching. And it doesn’t create real learning.”

### *Narrowing, Expanding, and Altering the Curriculum*

Another theme arising out of the interviews deals with the effect the AIMS has had on narrowing, expanding, and altering the curriculum. Even the same teacher sometimes held conflicting views about this concern. Teacher 1a saw narrowing as a good thing—saying it brings focus to “those skills and concepts that transfer to ten million other texts,” and she saw it as bad—“we spend so little time on narratives anymore, it’s just a small chunk of time!” At all three high schools, though, teachers

generally asserted that the AIMS had been affecting their curriculum in one of these three ways: narrowing, expanding, or altering.

At High School B, Teacher 3b introduced the idea of curriculum narrowing to the interview while expressing a concern that the English novel was being removed from the language arts curriculum and being replaced by more short stories or brief expository text similar to sections of reading on the AIMS test.

Well, my kids come up as sophomores, and they were talking about how they've never sat down and read a novel. And so, in the lower grades, they're teaching just little short blurbs; and that's great in terms of most of the reading comprehension questions are short blurbs, but they don't have the long-term reading abilities.

Teacher 4b concurred with this assessment by saying, "I don't know how many seniors told me [this year] at the end of *Frankenstein* that that was the first book they ever read from beginning to end. Seniors!"

Some of the teachers said that the AIMS had caused them to strip down their curriculum to a very basic level. Teacher 7b said,

I think for my class, we're really vocabulary based, even all the way down to literary terms because that's what's on the AIMS. I try to cover as many literary terms as possible because that's on the AIMS, and we're pretty much all my classes we're really basic. We went back to grammar. We went back to spelling. We went back to vocabulary. And I tell them, this is on the AIMS, this term here; and this is what it means, and we go through various examples through their writing, through their reading. So, my class is just basically grammar.

Four English teachers at High School B concurred that their curriculum had narrowed, mostly by way of reducing novels, creative writing, projects, poetry, anything not found in great detail or in high percentages of questions on the AIMS. Teacher 5b is representative of the group in her opinion: "You teach a novel, and it's just terms,

terminology, reading comprehension. And that's it; there's not projects, there's no plays, there's no skits. It's totally narrowed for us, and that's just reading, that's not even the writing portion."

The math teachers at High School B claimed that their curriculum had broadened, but saw it as a negative.

Teacher 1b: It's broadened. For example, they have a new PO that they set up that will be implemented next school year, which they just posted last month and that would tell you again we need to restructure. And they change a lot of things like ten percent. Ten percent is a lot because you still need to restructure all those things from first grade all the way up to twelfth grade. And they're still broadening it. They said, oh we're making it narrower for you. But technically, it isn't, it's broadening.

It is interesting that this math teacher saw broadening as a negative because they would have to restructure their vertical alignment and curriculum blocks across the school district at each grade level.

The teachers at High School A felt that the curriculum had expanded because of the AIMS and the structured, standards-based curriculum. These teachers suggested that before AIMS, there were very low expectations at this reservation school. They felt that the demands of AIMS, and their school's effort to implement tutoring and supplementary math classes, and set achievement goals had broadened the curriculum for students by forcing them to experience the full curriculum necessary to pass AIMS. It was an interesting take on the topic. Teacher 5a said,

I'm not sure this is correct, but I can see in a very high achieving affluent high school that [narrowing of the curriculum due to AIMS] than it does for us here. Because, in fact, this test, because of the things we've mentioned, the factors on students, on teachers, on what this test has done for us as a school, I think it expands the curriculum for a lot of the kids that we have here. Like, no, no, no, they can do it; they must do it. There's no way around it.

Immediately following this comment, a contradictory discussion ensued about the narrowing of the curriculum due to AIMS. The math teachers held that the curriculum had broadened, especially after the sophomore year in the higher level math classes, but the English teachers were conflicted. Teacher 1a talked about English teachers letting go of their personal interest units to focus on AIMS related content.

I think in the English department I've seen that happen a little bit where you know if somebody has a real favorite thing, a genre or whatever that they want to spend a lot of time on. And here's the plus and minus. For them, personally, shaving it down would be a negative, but for the good of the whole. It's probably a positive where they're not just one-tracked out you know. It just depends on the perspective on that.

Teacher 4a also seemed conflicted about whether it was a positive or a negative to “shave down” the English curriculum.

For me, I had a whole unit on poetry, and I talked to her [the curriculum leader] about this, and I was all excited. I was like, I'm gonna do this one week, and then all of a sudden I had, no, you gotta break that down; you gotta hit this, and we gotta hit this. You have flexibility, but then again you don't. So, you do have to shave it down, especially as English teachers. Sometimes you do things that you think the kids will love, and then you have to cut some of that out just because you gotta meet all the other standards. There is so much you can do with English, and it's sad we can only get bits and pieces out of it.

I asked the teachers at High School C if their curriculum had expanded or narrowed as a result of the AIMS. The math teachers insisted it had not narrowed.

Teacher 2c had the first and most concise response: “As a department, I'm positive it hasn't; they teach what they've always taught, and so do I.” Teacher 3c immediately agreed: “I also concur. [There are] some of the things that you have to hit harder on, but as far as dictating what I teach? No. The curriculum is taught, and they're there to get it. As it comes time to review [for the AIMS], you may hit it a little harder on particular

things that are gonna be there.” However, they did suggest that their curriculum shifts with the test each year, as they discover percentages of the standards represented in the test questions and identifying areas of student weakness. Teacher 3c touched on this as part of our curriculum narrowing discussion.

[On the AIMS test this year, he could] see there were a lot of questions on a couple of things that haven't been emphasized as much. And I'm wondering if that's really where we're headed and what's really the emphasis this year. And it was a lot on graphing of quadratic equations, and that hasn't been a hard emphasis and that's more of an algebra II versus an algebra I or geometry. But if we're gonna have ten questions of that, that's something that's gotta come back down.

Later in the interview, Teacher 2c noted a significant expansion in the curriculum, “There are certain things in the curriculum we have done. There's writing across the curriculum. We write in every math class and science and social studies.”

Later in the interview at High School C, I revisited the narrowing idea, but couched my query in different language. Instead of suggesting a narrowing (which immediately makes teachers defensive), I asked if anything in the curriculum gets altered as teachers analyze AIMS data and seek to make adjustments (the teachers at High School C said that this data analysis is an important part of staff meetings). Teacher 2c suggested there are changes taking place but was still somewhat reluctant to admit to it.

About all I can think of that we've done slightly different in math is the last couple of years, with pre-algebra, we cut our focus down a little bit. And there are certain areas that we've really focused on in math, in pre-algebra and basic graphing and basic equations; we've really beat those into them. We really want the kids to have a base in those, so they move on to algebra I. So, there's not as much depth that we've covered in pre-algebra the last two years. Me and the other teachers of pre-algebra, we've focused on certain things a little more, but we haven't got into any of the algebra stuff or the geometry stuff that is in the pre-algebra book. We haven't really gone there; we've focused on the algebra stuff a little more, uh, before they move on, but other than that, we haven't changed anything. We teach the standards.

High School C and High School A were the most defensive groups when it came to discussing possible changes in the curriculum due to AIMS. As the interviews went on, they became more comfortable sharing opinions, but at these two high schools, it seemed that the teachers were hesitant in their opinions. After the interviews and observations and document collecting, I could see that those two high schools had school-wide, even district-wide curriculum programs in place centered around AIMS scores, data analysis, and making adjustments in teaching or the curriculum. High School B had started down this path, but their school was not fully invested in the program and the teachers were much more open about their opposition to the possibility of completely entering into such a program.

At High School C, Teacher 5c was open about her curriculum adjustments.

See, I think I made a mistake because we had two years of high AIMS writing scores, so I thought, okay, I've been shortchanging reading because our reading scores are lower than our writing scores. I need to sit and emphasize reading more. So, I brought in young adult novels because I thought, we need to get them interested in reading, and if we tackle something that's easy to read, we can tackle it more in depth. But then you've got today's student who doesn't do homework, so I had to give up much more classroom time than I anticipated because they won't read at home. So, I had to devote class time to reading and analyzing reading in class, which meant less emphasis on writing. Did that translate into our low writing scores this year? I don't know. Our reading scores didn't go up measurably, so I'm thinking, okay, forget trying to get them interested in reading.

Toward the end of the interview at High School C, there was another hint from Teacher 2c that maybe the math curriculum was being affected.

Or, like number sense, one of the standards, that's maybe four questions on a [AIMS] test out of 85 questions. But if they don't know that then they're gonna struggle in some of the other areas. So there's some things I'd love to do, but I can't spend as much time as I'd like to do on them because I've got to get to the next standards.

At High School A, Teacher 5a responded to my query about possible narrowing or expanding of the curriculum. Her response was indicative of the general feeling the teachers at High School A and High School C seemed to have about the possible narrowing of the curriculum due to the AIMS. They generally thought the narrowing in on AIMS standards in the curriculum was a positive, at least in the lower high school grades. Teacher 5a said, “I think it has everything to do—it may as far as limiting the curriculum or whatever—with setting goals for us and our students to get to the point where we have the majority of tenth graders passing.”

#### *Cultural Challenges and the AIMS*

A theme arising out of the interviews at High School A and High School B, the two reservation schools, was the idea that the students’ success on the AIMS test might be affected by culturally specific challenges facing the Navajo high school students on the reservation. This finding resonated with my own experiences. I taught for two years and served as an administrator for one year at two high schools on the Navajo Reservation. Seeing a lack of research involving schools on the Navajo Reservation, and sensing there were unique challenges facing these schools with regard to AIMS testing, led me to include two reservation high schools in my dissertation research about AIMS effects.

The Navajo Nation is undergoing great changes as modern technology brings new connections to those who were once a relatively isolated people. Their love for family and traditions gives the Navajo people great strength. High school graduation rates increase each decade, and with the effective university and college outreach programs,

more and more adults are earning degrees and filling positions in the schools as teachers and administrators, and in the other prominent employment fields on the reservation: health care and government. There is much to be excited about on the reservation and the public school students are gaining strength each year. However, there are some unique cultural challenges facing these students that many educators feel the AIMS ignores as it acts as the gatekeeper to a high school diploma.

The Navajo Nation is largely rural desert. Though nearly two-hundred thousand Navajo people are estimated to live on the Navajo Reservation, there is no true internal economy. Most jobs are found in tribal, state, or federal government departments operating on the reservation, in local state school districts, or in various hospitals. Diné College is a small 2-year college with satellite campuses throughout the reservation, and four-year programs are available via various colleges and universities through extension programs; yet, the post-secondary education levels of the general population remain comparatively low. Many of the high school students are classified as ELL by the ADE because Navajo is usually a dual language, with English, in many of the Navajo households. The crime rate is high, with stronger gang influences than one might expect. Alcoholism is rampant, and drugs are a major problem throughout the reservation. Estimates vary, but it is widely suggested that well over half of the adult population is unemployed due to the lack of a local economy.

A high percentage of the students live in scattered homes accessible only via dirt roads that often become impassable in snowstorms, spring snow thaws, and significant rainstorms. There are very few “neighborhoods,” and the few often do not have



sidewalks, parks, and curbed streets as is common in modern American cities. There are no “big box” stores on the reservation. The entire nation travels off the reservation to “border cities” like Gallup, New Mexico; Farmington, New Mexico; Page, Arizona; Flagstaff, Arizona; or Winslow, Arizona where Wal-Marts and shopping malls can be found. The capital city of the Navajo Nation, possibly its largest city, has well under 4,000 residents, no movie theaters, no big box stores, no grass parks, no public swimming pool, no shopping mall, and only one national chain, family-style restaurant: a Denny’s. The students living on the reservation are connected to the outside world via the recent proliferation of satellite television and occasional trips off the reservation to the border cities for shopping or youth sporting events. Most of the population lives in scattered modular homes or single and doublewide trailers, many of which require water to be trucked in from scattered wells to holding tanks at each home site.

School absenteeism is a major problem facing teachers and administrators, and transferring from school to school is a frequent occurrence for many students. A positive aspect of the culture that affects school attendance are the traditional Navajo religious ceremonies and the Navajo people’s commitment to family and clans. Religious ceremonies are performed for a variety of reasons, especially healing, and often require students to be out of school for days, and at times, weeks. Despite all of these unique challenges, the thousands of Navajo Nation youth are required to pass the AIMS for high school graduation with very little in the way of added assistance to the reservation high schools for academic programs or remedial interventions. The teachers at High School A and High School B discussed many of these challenges with regard to the AIMS.

Teacher 3b: There are a lot of cultural things, too. We've read some of the AIMS test materials and major locations, you know, major cultural concepts that are not familiar to the Navajo. And the kids didn't understand those things. They may have understood what an analogy was, what a theme is, but because they didn't understand the cultural reference, they missed the question. So, it wasn't a very good judge of whether or not they could actually perform the task for the standard that was trying to be tested.

In the middle of a longer conversation on this topic at High School B, there were several comments indicative of the ideas being discussed about how the Navajo student often has a disconnect with seemingly innocent cultural references. For example, there are windmills scattered throughout the reservation that pump water into holding tanks for people and livestock. Many people use these holding tanks for various purposes. Teacher 5b said, "I had a kid who didn't know what a propeller was but knew what a windmill was." Teacher 7b said, "I had a kid who didn't know what a flowerbed was. Out here, who has a flowerbed? We don't have flowerbeds." Teacher 6b shared an example of a biased math question: "For example, [a] question in the math test... something like [someone] lives six blocks from Frank and three blocks over there.... There's no blocks here!" A similar conversation at High School A brought up parallels to the High School B interview. Teacher 3a said, "I've seen words in the practice test like recreation center. What in the world! This just doesn't link to our children's lives at all. A lot of them haven't been at places like a recreation center." However, Teacher 6b did acknowledge efforts by the ADE to eliminate this sort of testing bias: "But we did have one of our faculty go down and address cultural bias on the AIMS test, so I'm just saying they're aware of it. I'm not saying it's totally eliminated, but they're at least aware of it."

Navajo Reservation school districts often experience high teacher and

administrative turnover rates; a primary reason for this is, non-Navajo (the majority of the school staff usually) employees typically live in district housing on school property. Most teachers and administrators are just in a district for a few years, and then they move on. At High School B, they have had difficulty, in recent years, retaining a full staff of counselors, and they experience considerable student body turnover year-to-year and even semester-to-semester. As a result, there were frequent major schedule changes for a high percentage of the student body. Teacher 8b discussed this.

And at this school, in the English department, it's hard to help kids advance in their learning and advance for their AIMS scores and try to get all those skills in when, in this school, our kids are constantly being shuffled. It's not just the first week of the school year; they get shuffled from class to class throughout the entire semester, throughout the entire school year. And they don't necessarily know what's coming. My kids, all of a sudden, they're handed a schedule change and my kid has been moved from my class to [another teacher's] class a week before the quarter change. And how is that gonna help the kid prepare for AIMS?

I heard the argument that in a vertically aligned, blocked curriculum district, this should not matter—the transitions from one class to another should be smooth. However, as teachers pointed out, the long-term subs, high teacher turnover, and the non-standard nature of a Navajo student's educational experience on the reservation made these schedule changes significant.

Teachers at High School B pointed out that poverty, the isolation of the reservation, exposure to English, crime, low parental involvement, and drugs and alcohol are powerful challenges to their students. As these two reservation schools seek to implement AIMS-driven, vertically aligned, blocked curriculum programs, they would need years to see a thorough effect among their unique student body; yet, the percentage of students who pass AIMS must rise each year per NCLB requirements toward the

future goal of 100% proficiency. With students in their classrooms facing such unique challenges, teachers at these two schools seemed under an enormous amount of pressure “to make sure they are at par” with their statewide peers. Here are some excerpts from their conversations.

Teacher 1b: We’re in a poverty stricken area, so our kids are dealing with a lot of problems before they go to school, and then when they’re in school, its like you’re telling them, okay, you need to do this; you’re supposed to be a regular student. It’s like we’re just concentrating on school. Studies show that since kindergarten, they’ve always been lower than the standards, that you always have to pull them up so they meet the standards. It’s kind of hard. Those things really affect us as teachers here, and the pressure goes down to us to make sure they are at par.

Teacher 3b: In terms of English, [our kids] have far less exposure to the English language overall...and that is absolutely portrayed in the AIMS test. I mean, they just don’t have the same exposure in terms of even the little things like reading material availability. They don’t have that here. You don’t see kids with libraries at home, or there’s one bookstore 26 miles away that’s in a mall that’s overpriced. They don’t have the money for books; they don’t have the money for gas to get there and get back. So, they don’t have the same exposure to the English language that kids off the reservation do, and that’s portrayed in the AIMS test, and there’s no compensation for that.

Teacher 3b: They’re not enjoying the extracurricular childhood activities of middle school and high school, so they do have less connection with the school, you know. You see that in terms of parent involvement, too. Parents feel less connected to the schools than off the reservation schools.

Teacher 5b: I think in the end, I know there are a lot of studies that say all of this, that say that Native Americans have high alcoholism, have high domestic violence, have high poverty, have all these social issues connected directly to Native Americans, not just here on our reservation. It’s all the reservations across the country. I mean, my God, you know we had several suicides this year [this high school experienced three in the semester during which this interview took place].

Teacher 4b: I think part of the schema problem [for understanding AIMS content] is the isolation of the reservation. They’re focused on basic survival.

At the close of this conversation on cultural issues and AIMS, this exchange took place

between several teachers:

Teacher 3b: As an educational institution, we're focusing on the wrong thing.

Teacher 5b: Yeah.

Teacher 3b: We're focusing on the AIMS.

Teacher 7b: We don't have the mainstream middle America students here. We have students who are totally different.

Teacher 3b: We're not suburban white America. I mean, these kids, I can't compartmentalize stuff when something happens at my house with something that's going on at school. I mean, it's hard for me to focus when I'm upset, and we expect our kids to completely compartmentalize it—you know, I am sorry everything at home is terrible, but you need to be happy and ready to study when you get to school; that doesn't work.

The feelings about cultural challenges and the AIMS were far stronger and the conversation lasted much longer at High School B than at High School A. I never could figure out why. An interesting comment coming out of High School A dealt with the perceived equalizing factor that the AIMS had despite the cultural challenges facing the Navajo students.

Teacher 3a: It makes passing the AIMS such an accomplishment because it's not just, oh I passed, now I get a diploma. But it's, hey, I did what those other kids are doing who have a lot more economic and you know educational resources available to them that our students often don't. I think it has created this neat sense of pride among the students when they pass it, beyond just the diploma aspect, but the accomplishment side.... I'm proud of them because they are taking the exact same test as people in very different circumstances.

While the teachers at High School A took pride in the standardization of their curriculum and, to some extent, their students, the teachers at High School B were frustrated by this standardization. Talking to the teachers at both schools, one can sense their concern for the children in their classrooms, and their deep worry for them as they face challenges

that most outside the reservation cannot comprehend. The possibilities for helping these students are endless, yet the reservation high schools remain under-funded and largely ignored in the AIMS discussions as these students deal with the pressure of trying to rise up to the standards of the AIMS while living in unique circumstances.

### Classroom Observations and Document Collection

Following the interviews at each school site, I observed one class for each teacher involved in the focus groups and, after each observation, sat down to discuss the observations and the interviews with most of these teachers in their individual classrooms. I also collected documents that teachers felt represented AIMS influences, or a lack thereof. The classroom observations at High School B took place during May 2008. The classroom observations at High Schools A and C took place during September 2008. I conducted five observations at High School A, eight observations at High School B, and six observations at High School C for a total of 19 classroom observations. The classes were generally 50-90 minutes in length. I spoke with each teacher privately, typically during their prep hour, to further discuss ideas from the interviews and the classes I observed. Most of my documents were collected during these visits and some of the “prep hour discussions” involved documents teachers shared with me. I took detailed notes about what I saw during the observations and about the conversations between classes. I reviewed and evaluated these observation notes and the collected documents to determine correlation between the site visits, collected artifacts, and the emergent themes and ideas from the focus group interviews.

The classroom observations and collected documents, and the conversations that took place in the classrooms, paralleled the focus group interviews. Those teachers who suggested the AIMS was influencing their curriculum and instruction had classrooms and lesson plans that reflected the influence. I saw ADE standards posted on the walls, verbatim, in a number of classes. I heard frequent mention of the AIMS during lessons, and when I asked these teachers if the “AIMS talk” was because I was there, they said that they frequently refer to the AIMS during lessons—one teacher said this was to “motivate” the students and give them “a reason for learning.” I saw a strong influence on every English classroom with posted Six Traits for writing, and a heavy prescription of five-paragraph essays. I saw some teachers pushing vocabulary because they felt it would help on the AIMS. Another teacher was reviewing a nonfiction reading assignment, which she created in response to the previous years’ low AIMS scores in the non-fiction areas of the reading test. High School A and High School B gave me curriculum “maps” and documentation on the vertical alignment and blocking efforts; High School A shared documents from their staff development where they analyzed AIMS test data.

The conversations with the teachers in their classrooms paralleled the interviews as well. Not a single teacher retracted their focus group comments; most of them reiterated or elaborated on their views. As noted earlier, for those teachers who shared details about how their class “runs” now, I saw these classes operating exactly as described. I know my observations were only a brief glance into each school, but I was surprised that I did not see a single lesson that might be called “fun” or “creative.” Every lesson was essentially following a typical format: teacher presents, students practice,

teacher revisits concepts, students practice. There were no hands-on activities, projects, cooperative learning, or nontext centered lessons observed.

One of the more interesting documents shared with me came from High School A. Hanging on the teachers' walls in several of the classrooms was a form that said, "School-wide Priorities." I was informed that teachers had created the document under direction of the administration during the staff development days prior to the beginning of that school year and that most, if not all, teachers had this on their walls as a philosophical guide of some kind. On the left side was a column titled "Priority," which read:

- SEI-style instruction [SEI stands for Structured English Immersion]
- Stronger Vertical Alignment
- AIMS Prep—Tutoring, advisement, life skills class, academic intervention in all classes.

On the right side was a column titled, "Resources Needed," which read:

- Professional development-level planning
- Time, guidance, commitment from more team members, curriculum days
- Professional development on AIMS-style reading, resource list, practice resources, 6-Trait rubric and training, AIMS test committee, AIMS Prep curriculum, teacher resources.

Seeing this document and the documents showing the intense push to analyze data and reform curriculum and instruction based on this analysis, I could see the foundation for the responses the teachers from High School A had given in the interviews.

Another interesting document I received came from High School B. This was a printed and bound copy of their curriculum map—standards and POs broken down by



grade, class, and quarter. Aligned to this guide were their benchmark tests for each quarter, which were supposed to cover the POs and standards for that quarter. The notable thing about this document, though, was that fact that the district had renamed Algebra I and Geometry in the guide: Algebra I had been renamed, AIMS I; Geometry had been renamed, AIMS II (and this name change was also reflected in other school documents, like class schedules).

Ideally, I would have liked to observe more classes, but I am confident that the observations conducted were adequate for my purposes. It was quickly evident that the interview responses were reflective of what was happening in these schools. The documents collected supported the conclusion that each school had reformed their curriculum around state standards and that these efforts were, to some extent, affecting most of the classrooms. Only two teachers, and both at one school, held to the position that the AIMS was having no effect on their curriculum and instruction, though observing their classes revealed little variation in the teaching methodology so prevalent in all other AIMS-centered classes.

In reviewing my notes from the observations and document analysis, no sharp contrasts between information gathered during school visits and data from the focus group interviews emerged. By this, I mean that observations and data analysis fit comfortably within the context the teachers had created in their interviews by describing their schools and classrooms. What I saw in the field was consistent with the descriptions of influence, change, or practice that the teachers described in the interviews. The strongest theme that did rise up out of my observation notes was the ease with which I

could observe the cultures of testing within the schools. This culture was pervasive in every classroom. The pedagogy described in the interviews and ascribed to AIMS effects was indeed taking place. No maps, curriculum overviews, lesson plans, work on the walls, portfolios, or anything observed could be characterized as cooperative learning, project-based learning, hands-on learning, or any other principles of nontext centered student discovery type learning steeped in critical inquiry and higher-level thinking. The school and classroom cultures were very test-oriented, teacher-centered, and text-based, and the curricula adhered strictly to the standards leading toward the AIMS.

At the conclusion of this presentation of the data collected, I want to touch on several points before analyzing the data in the concluding chapter. First, I did not notice any connections between years of experience, gender, or ethnicity and patterns in the responses. In fact, responses seemed tied to one consistent factor: the school site and the philosophy guiding the education system within that school at the time of the interview. Second, after talking to the teachers in the focus groups, revisiting them in their classrooms in private conversation, and observing their teaching, I believe each of these teachers was honestly expressing his or her genuine perceptions and each had a real concern for the success of his or her students. I did not walk away thinking, “Boy, I wish so-and-so hadn’t volunteered.” Did I see what might be called “poor” teaching on occasion? Of course I did. Did I see excellent teaching? Yes. Did I see both veteran and what we might call “beginning” teachers? Yes. However, I did not sense any weakness in the body of participants as a whole. Their varied experiences, views, and abilities enhanced the overall body of data.

Finally, I entered this study desirous of achieving a meaningful more insight through an exchange between educators in light of Eisner's work on educational criticism. Eisner wrote, "what teachers and students do is influenced by their location in a system" (1998b, p. 2). I saw both myself and the teachers as important participants within the systems of public education seeking to more fully understand these systems of influence. I felt that in connecting the multiple lines of experience and expertise under the umbrella of one study into the perceptions of practicing professionals, an important qualitative discovery might be made about preconceptions I had as a researcher and suggestions about testing effects arising out of the literature on high-stakes testing. I feel that discoveries have been made through this study. I am confident that those of you reading this chapter have already arrived at some of your own conclusions, raised your own questions, and possibly formulated some of your own implications for possible future research. The next and concluding chapter, summarizes my analysis of this data, provides my interpretations as an educational critic, and offers my suggestions for future potential research on the issues raised in this study.

## CHAPTER V

### ANALYSIS, INTERPRETATION, AND IMPLICATIONS

As suggested at the beginning of Chapter IV, I believe that some of the analysis and interpretation of the data came naturally through the narrative presentation in Chapter IV. Threading my thoughts and ideas about what was being said through the interviews provided at least an introductory level of analysis. However, there is need for elaboration here. Looking at each of the six emergent themes, I have several observations.

#### Observations

“Creativity in the classroom” was a theme that arose out of strong responses from High School B and High School C, though there were some interesting contributions from High School A. Teachers at High School B and High School C felt strongly that AIMS was adversely affecting their creativity in the classroom. My impression from the teachers was that while AIMS might be a good indicator of basic skills and knowledge, they were concerned that it was the cause of many teachers adjusting teaching, assignments, and assessment to match the format of the AIMS. They were dropping untested standards like presentation or speech-related standards, and they were narrowing their math, reading, and writing to reflect AIMS-style math problems, AIMS-style reading (typically short, nonfiction pieces), and AIMS-style writing (five-paragraph type essays). This has serious implications. Both math and English teachers suggested that their teaching had begun to reflect the multiple choice, basic skills focus of the AIMS.

Among the English teachers, the most repeated example of this shift was the Six

Traits guided, five-paragraph essay. This is the style of writing assessed by the AIMS, and consequently, English teachers at all three schools seemed to be saying this had become their primary mode of writing in their classrooms. In fact, I heard repeated references to anything other than the five-paragraph essay as “creative writing.” The interesting thing about this, though, was the fact that these teachers had seemingly come to define creative writing as literary fiction or poetry—something teachers at High School A said most students just do not care to do. One of the possible consequences of teachers replicating AIMS-style writing in their classrooms is the narrowing of the definition for any other style of writing. “Creative writing” can be more than just literary fiction; it can be many things: a page for a technical manual; translating a formal letter into a less formal email; creating ten rules to govern an organization; rewriting a school policy; writing an instruction on how to complete a math problem; writing an incident report for work or school; the possibilities are endless. True creative writing accesses higher-level thinking and broadens a student’s writing skills set to better prepare them for the real world application of writing. AIMS had seemingly narrowed several of these English teachers’ focus in these three schools to Six Traits guided, five-paragraph essay styles of writing and inadvertently narrowed their definition of creative writing.

This brings us to several important concluding questions on this matter: Is such a narrowing taking place? Are classrooms statewide truly beginning to reflect AIMS-style math, reading, and writing? If so, is this a good thing? Are English teachers narrowing their definition of writing to reflect the five-paragraph essay? Is the lack of more “creative writing” in English classrooms pervasive statewide, and if so, what are its

effects? These are important questions arising out of this section of the data that should be investigated by future research.

The themes “vertical alignment and structured curriculum” and “standardized curriculum versus nonstandard students” combined to form one strong implication. Teacher 3a said, “Their education shouldn’t be primarily based on the teacher’s likes, dislikes, and personality.” There is some truth to this statement. Although, some would argue that removing teachers’ likes, dislikes, and personality creates a standardized, bland, cookie-cutter set of classes within a school and removes creativity and innovation and diversity from the classroom curriculum and instruction. The idea, though, is that we do want to eliminate poor teaching that does hurt students academically when, for example, a teacher in love with poetry spends 30% of the school year in a sophomore English class reading and writing poems (an extreme example, but indicative of the line of thinking). But it is Teacher 3a’s comment following this quote that has the more serious implication: “[A child’s education] should be based on the content that’s appropriate for that grade level, for that course.”

My general sense from the teachers at all three schools was that the vertically aligned, blocked, or mapped curricula had caused a shift away from students as individual learners in need of a dynamic curriculum toward an AIMS-driven static curriculum focused on prescriptive benchmarks for grade-level appropriate learning and assessment. This does not imply that these teachers cared about their students any less. Throughout the interviews and in my classroom observations, I could see that all of these teachers truly cared for their students and were concerned about their success in school

and general well-being. I do believe, though, that the systems of education these teachers were describing in their schools, and the frustrations they expressed as reflected in Chapter IV, were indicative of a shift from student-centered education to system-centered education.

While many educators espouse system-centered education as an equalizing force in education necessary for a meritocratic education system (an idea at least one teacher at High School A touched on regarding her Native American students “accomplishing” the same things as “other” students across the state), there are many educators who see the possible disconnect between standards-based education systems and educating individual children as a cause for concern. The implication here is that we investigate this matter specifically. Is the system-centered education at these three schools pervasive in Arizona, and if so, is it causing any kind of detrimental loss of focus on individual students and educating the whole child? In addition, regarding the Navajo Reservation, is this system-centered education really in the best interest of Native American students?

The primary cause of concern I have arising out of this study is the idea that the participating teachers in this study are genuinely concerned about their students—they really do care about them and want them to learn and be successful—but the AIMS seems to be dictating an important shift in the philosophy which shapes their practice. The question guiding their teaching is no longer, “What do my students need to know to be successful in life, and what does this student in particular need from me as a teacher today?” Rather, the guiding question seems to be, “What do my students need to know to be successful on the AIMS, and what does the curriculum map tell me I need to teach

them this year or this month to make that happen?” I think this is why some of the teachers expressed confusion and frustration as they discussed these issues. They believed they were being good teachers by finding creative ways to prepare their students for the AIMS and to pass the high-stakes exam. Yet, they seemed confused about whether this is really what they wanted to be doing. They seemed frustrated by the situation such as that in which the system had placed them. They seemed to be asking, “Am I doing the right thing here?” The frustrating thing for me as an educator and researcher is my response has to be, based on the paucity of data, “We don’t know, but smile, hopefully, you’re doing the right thing.” Perhaps the future research question coming out of these two themes is, “Does preparing a student for the AIMS in a vertically aligned, blocked and mapped curriculum system meet the needs of individual students and properly prepare them for life?”

Considering these concerns in light of the epistemological underpinnings of this study, we might consider other questions based on the theoretical framework of this research. Would Dewey or any experienced educator claim that preparation for one test (or even several) meets the needs of individual students? Is the only knowledge that counts, knowledge that is grounded in data and assessed in standardized forms? Would these teachers in this study, who seemed to repeatedly return to the idea that the blocked and mapped curriculum caused them to feel a pressure to move on or to teach only a narrow thread of knowledge, say that they feel this is the best way to prepare individual students for life and society? An accumulating body of literature, from philosophers like Dewey to contemporary educational philosophers and practitioners, plus a growing body



of qualitative studies, all seems to provide credible reasons to pause and consider the propriety of standardized approaches to education when taken to the extreme and so focused on high-stakes exams.

The important implication rising out of the section “reviewing or cramming for the test” is summed up in this phrase from Teacher 5b: “Graded by the state department.” In many ways, this might be the grand implication arising out of this entire study. Schools, teachers, and students are all being “graded by the state department,” and these teachers seemed to be trying to make sense of this reality. In the world of public education, the state departments of education are typically distant and somewhat mysterious to the average school administrator and classroom teachers. As reviewed in Chapter I and the brief history of standardized testing, state departments of education have been handed a great deal of power.

One of the more interesting documents I collected in my research were the ADE’s School Report Cards for each of the three schools in this study. Each year, every school in Arizona receives a report card that notifies the public how that school is doing. “What is on that report card?” you might ask. First, prominently headlining the report card is the school principal’s full name and personal school email adjacent to the school’s address and contact information. The school’s mission and goals follow this information along with the number of students served and grade levels. Next, the AZ Learns profile for the school tells the public whether the state feels the school is Performing or needs improvement of some kind. There is also a line designating the Federal School Improvement Status. These are followed by the AYP Status (met or not met).

Prominently centered on the report card, and the focus of the entire report, are the Test Results for the AIMS and state TerraNova exams. Scores are indicated for the past 3 years based on percentage of students passing in the three testing sections: Math, Reading, and Writing. At the bottom of the page are a variety of reported statistics: on campus incidents, number of ELL students, attendance rates, promotion rates, drop out rates, four-year graduation rates, and 5-year graduation rates.

As noted, the AIMS scores are the focus of the report and play the central role in determining a school's status with the state and federal governments. The scores are presented in color-coded bar graphs and show the school's rise and fall year to year in percentage of students passing the AIMS. Reading these report cards, it is easy to see why these teachers had come to define their profession, their classroom experience as teachers, by student success on test scores.

The ADE, quite literally, as evidenced by their annual School Report Cards, makes the statement "graded by the state department" a powerful truism. A principal whose name is printed across the top of that report card with a personal email is likely going to read those grades with a purposeful eye. Teachers are likely to be held accountable for those grades, and the students of these teachers will indeed be held accountable—no passing AIMS score means at least some kind of remediation and, at worst, no graduation. No matter how you look at it, being "graded by the state department" is going to affect a school and its teachers. Hence, much of what schools, teachers, and students do is "for the test." The implication arising out of this study is that this grading effect is potent and needs further inquiry as to its outcomes and their

connections to students.

In the section “narrowing, expanding, and altering the curriculum,” an interesting fear rose to the surface of the discussions. In education, “narrowing” causes an almost immediate reaction among teachers when discussing their curriculum. Most of the teachers immediately became defensive when I broached the subject using the word “narrowing.” As data in the other sections or themes shows, various types of narrowing was taking place at all three schools in multiple classrooms. I sensed that the word “narrowing” had become negatively associated with AIMS or high-stakes testing in general, and that these teachers were reluctant to admit to any kind of trimming down of the curriculum, even if it might be a positive trim as some teachers suggested. Focusing on what is important can be a good thing, after all.

In retrospect, I wish I had pushed this further, perhaps by revisiting the topic while talking to teachers in the comfort of their own classrooms. I wonder what fears teachers have as they face the shifts in their practice due to AIMS. I did not investigate this idea. It would make an interesting research study. Finding out how these fears affect teachers would stimulate the general discussion on high-stakes testing. Are teachers motivated? Depressed? What good arises out of these fears? What bad? Is fear a natural and necessary aspect of teaching in a high-stakes environment? Expanding on this theme, and perhaps more important than investigating teachers’ fears, I would like to know what fears students have in such a testing environment and how this might affect them for good or bad.

After reviewing the data under the theme “cultural challenges and the AIMS,” the

first question that comes to my mind for future research is: Do the unique set of dynamics (teacher and administrative turnover, frequent shifts in student population, lack of stability within a school, poverty, crime rates, language barriers, and so forth) play a significant role in affecting AIMS results, and if so, should the state acknowledge and address this issue specifically among Native American student populations if the AIMS is required for graduation from high school? These teachers suggested that there were unique challenges facing their students, and they were concerned with their ability to help students meet those challenges regarding passing the AIMS test. Related to this, and along the same lines as a question posed earlier, researchers should probably investigate the following question: Is the AIMS causing teachers to focus less on the unique knowledge and needs of Native American children sitting in front of them and focusing instead on producing test scores that meet state and federal requirements? Does the cultural mismatch between the AIMS tests and the Native American students' background knowledge from lived experiences create an unfair barrier to the already difficult-to-achieve standardized expectations of these tests?

After analyzing the data and offering these interpretations and implications for future research, we come to an important conclusion: we need more research. We need to know if these effects arising out of this study are widespread, and if so, are they in the best interest of the students. My sense is, based on the interpreted data from this study, and my own experiential understanding, that the shifts taking place are not in the best interest of our students. The shifts seem philosophically opposed to assertions grounded in Dewey's democratic progressivism and many current learning theories that espouse a

focus on the individual student as a child living life. I worry that our classrooms are drifting from the concept of helping children to reach their full potential as human beings, and instead are focusing on helping our children reach their full potential to score well on a high-stakes exam.

The classroom observations, the documents collected, and the ideas expressed by the teachers in the interviews develop an introductory picture of the culture of testing the AIMS seems to be influencing in these schools. As the philosophy of education in these schools becomes centered on a testing system rather than the students, the pedagogy seems to be reflecting the shift. While some might argue that this shift is good for the students, it requires an investigation into the definitions such cultures of testing give to words and phrases like “challenging students to achieve more,” “best practices,” “success in school,” “sound pedagogy,” and the purposes of schooling. Educators need to evaluate their schools, determine if a culture of testing has developed, and have meaningful conversations with all stakeholders about the consequences and meaning of allowing such a culture to develop and exist within a district or a school.

I believe that schools and districts can focus curriculum and instruction on principles of child centeredness as part of Dewey’s vision of democratic progressivism. Critical inquiry, higher-level thinking, and experience-based education can be central to the day-to-day instruction in our classrooms as part of a school’s effort to improve student learning. Arguably, this can even have a positive effect on test scores if tests are sound and make direct connections with important aspects of child-centered learning. Future studies might sharpen our understanding by investigating systems of education

which are student centered rather than system centered or by finding and studying teachers within test-centric environments who successfully adhere to principles of Dewey's democratic progressivism even within a system that does not focus on the child.

It is my belief that cultures of testing exist in schools and districts without investing time and resources in properly determining the consequences of such shifts on students and teachers. I believe the curriculum and instruction begins to change, in ways that contrast sharply with Dewey's vision of democratic progressivism, and these changes often go undetected because districts are busy analyzing test scores and adjusting practice to influence scores. Districts embracing cultures of testing may unwittingly allow the definitions of student success and student achievement to be defined by test scores without determining the consequences of such definitions on pedagogy and student experiences in school, .What is the likelihood of nurturing Dewey's ideal of curriculum which focuses on experience based on an individual child's needs, tying together both what she knows with what she needs to know to be part of a larger society in such a test-centric environment? I believe meaningful conversations should take place between stakeholders, beginning with school administrators and teachers doing as I have done here, using Eisner's educational criticism, to describe, interpret, and evaluate individual schools and classrooms. With such data in hand, these stakeholders can then make important decisions about what is best for their students based on a clearer understanding of the consequences of embracing a culture of testing. Armed with carefully constructed rationale for their decisions, they can communicate with policy makers more effectively, working strategically for what they believe to be in their students' best interests.

## Conclusion

Eisner (1998a) made the following statement about the relationship between researcher and teacher.

The relationship between researcher and teacher...is one of mutual inquiry and negotiation. Their generalizing qualities are not so much located in Truth, as in their ability to refine perception and to deepen conversation.... Such a conception of generalization lightens the burden. This lightened burden is not to be regarded as an invitation to irresponsible description, interpretation, or evaluation, but rather as a reflection of the recognition that generalizations are tools with which we work and are to be shaped in context. They are part of the substantive exchange between professionals with their own expertise, not prescriptions from the doctor. (p. 205)

Approaching this study as an educational critic in the mold of Eisner, I was able to see that the curriculum is being narrowed by standardized testing and that the creativity and artistry of teaching is being lost in these schools. Dewey once wrote, "Education...is a process of living and not a preparation for future living" (Bruner, 1966, p. 211). In this light, our American public education system is focusing less on education as a part of the life of a child in the midst of living, and drifted into a view of the child as a part of the manufacturing system guided and monitored by high-stakes testing designed to successfully pass percentages of students to the next level. This study has been an attempt to develop a clearer understanding of what is actually happening to curriculum and instruction in several public school classrooms as a result of one state's high-stakes exam. It was an opportunity to listen to professionals whose voices are too often marginalized, and in so doing, refine perceptions, and provide a picture that will deepen the national and state conversations about high-stakes testing effects.

At the end of this study, I have three main conclusions, three primary ideas that

emerge out of my consideration of what these teachers had to say, what I was able to observe, and what I learned through the analysis of relevant documents. First, the AIMS is having tremendous influence on the curriculum and instruction in these teachers' classrooms, particularly because of school-wide, and even district-wide, curriculum reform efforts to create standardized programs of instruction centered on ADE standards, vertical alignment, and principles of curriculum blocking and curriculum mapping that move students programmatically toward the culminating AIMS exam. Second, I sensed a great deal of conflict within these teachers as they listened to their own voices in the discussions and, perhaps, heard their own ideas for the first time in such a formal setting. Many of them seemed torn between supporting their school's curriculum program and the excitement that can come with seeing percentages of students improving on a standardized test, and their frustration with the significant effect it was having on the transformation of their profession. Teacher 4a expressed this internal conflict when I asked if the pleasure of teaching had shifted one way or another since the AIMS.

Mine is both. You know like the days when you're just like, God! I gotta meet those standards! Gosh, I gotta get these students prepared you know. Then, you're just like, I hate teaching! I hate teaching! Get me a beer! Sorry. I mean, you know, you just get tired of it. You get fed up with it. I'm not data driven. I hate looking at data, but I have to for me to be able to be a good teacher. It's just some people have the eye, and I just don't. And then there's just days when my kid can write a poem [and I say], Oh my gosh, you wrote this? This is beautiful, you know. I've met some standards; then I feel happy. I feel that pleasure. I love it. That's when I want to come back the next day. And then there's days when I just want to shoot myself.

My third conclusion is best expressed in this line from the above quote: "I hate looking at data, but I have to for me to be able to be a good teacher." This philosophy seemed pervasive at the three high schools. Part of the internal conflict I sensed in



teachers was their reaction to the realization that this philosophy had come to define their school and in some ways their classrooms. Repeatedly, I heard teachers defining success in the classroom and pleasure in teaching according to helping students improve on the AIMS test. Two teachers at High School A made two lengthy comments about the pleasure and “rush” they felt in increasing test scores and helping students pass the AIMS.

Teacher 2a: This data, I mean, we don't have to look at as a whole; we can look at individual students as well because we see some students who come into our classroom their math scores are so low. I mean just individual students. Then at the end of the year, it's way up here. I mean just that individual student—the fact that he or she learned so much. It really makes you happy. For example, this year I had a student who came into my classroom. He was so negative. He didn't want to learn. He told me himself, I'm just here because I'm here. I'm not here to learn anything. It's what he said himself. And at the end of the year, he found that he passed his AIMS. And then he gave me hugs and said, oh you did so much for me. I mean these are the types of things that really make me want to continue teaching, just even if one student says that to me, that's a big thing to me. But as a whole, yeah our scores yeah they may increase slightly, but still those individual students, those are the ones that really make a difference in my teaching and what I feel about my teaching.

Teacher 3a responded to this comment and took the idea of AIMS defining happiness even further.

And I think the very first year that you teach—you're AIMS driven—the AIMS I think causes more misery than pleasure. Because you don't have any results from any of the students to see what you were able to help them accomplish. But once you get those results and you see the impact of what you did, and as teachers you look at your students results and you can say, oh not only did they pass or not, but you know my students overall, did very well in this area. It's an affirmation that drives you, you know what I mean, to continue doing what you do and feeling good about it the following year. Does that make sense? And also, it is getting to those individual students looking at the individual student's scores. You know we're supposed to care about the big picture, because it's AYP, but the big picture isn't what excites me. It's not what makes me shout and it's not what makes me cry. When I see those scores, it's that child that has worked so hard and just didn't make it. I want to sit and cry because I know they are sitting and crying when they

get it. But then the ones that went from Falls Far Below to Exceeds, even though they have always fallen far below, that last time they're to Meets, you know, where they finally pass. You're not supposed to jump two categories in one test taking, but the kids who do, oh it's a rush! And although you don't get the data every day, that affects you every day, but on the low days, that's something you can mentally revisit is you know students are able to accomplish something because of experiences I can provide in my classroom. Does that make sense? It's kind of the warm - you know you're supposed to have a rainy day notebook that you go back through and it's all the warm fuzzy notes that you read when you're really feeling down. Like it's those individual students and what they accomplished on the AIMS that is my mental rainy day notebook.

Among these English and math teachers, it seemed that a large part of their existence as teachers had come to be defined by the AIMS test. Each school was so fully invested in curriculum programs centered on mapped standards that moved students toward the AIMS that some teachers had even come to define their satisfaction in teaching based on test scores.

These three conclusions correlate with some of my own preconceptions about the effects AIMS might be having on the curriculum and instruction in these schools. However, my personal conclusion is not, "Aha! I found what I was looking for!" Rather, I feel that a number of questions have arisen out of this study that require further investigation and discussion. First, how pervasive in Arizona is this curriculum reform that centers on vertically aligned, blocked, and mapped curriculum? Second, where we do find this type of curriculum, is it really what is best for our children? Third, does this type of curriculum ignore the unique challenges facing nonstandard students plugged into a standardized curriculum program, particularly Native American students? Fourth, how widespread are the conflicted feelings expressed by the teachers in this study? Fifth, what type of influence is this type of curriculum having on the practice of teaching and

teachers' feelings towards their profession? Finally, do the systems of education described by the teachers at High School A and High School B best meet the needs of the Native American students these schools seek to serve? I believe these questions need to be examined in the state of Arizona. The AIMS appears to be having a dramatic, and I believe a negative, effect on the three high schools in this study, and opening future studies up to a more comprehensive evaluation of the topic might provide additional compelling data for a deeper discussion of the issues raised here that can contribute to increasingly defensible approaches to reform in our public schools.

## REFERENCES

- Abrams, L.M., Pedulla, J.J., & Madaus, G.F. (2003). Views from the classroom: Teachers' opinions of statewide testing programs. *Theory into Practice*, 42(1), 18-29.
- Amrein, A.L., & Berliner, D.C. (2002). *High-stakes testing, uncertainty, and student learning*. Retrieved May 10, 2007, from <http://epaa.asu.edu/epaa/v10n18/>
- Au, W. (2007). High-stakes testing and curricular control: A qualitative metasynthesis. *Educational Researcher*, 36(5), 258-267.
- Baines, L.A., & Stanley, G.A. (2004). High-stakes hustle: Public schools and the new billion dollar accountability. *The Educational Forum*, 69(1), 8-15.
- Ball, D. (2001). *How to do your own focus groups: A guide for trial attorneys*. Louisville, CO: National Institute for Trial Advocacy.
- Bangura, A.K. (1992). *The limitations of survey research methods in assessing the problem of minority student retention in higher education: The focus-group method as one alternative*. San Francisco: Mellen Research University Press.
- Beck, F.D., & Shoffstall, G.W. (2005). How do rural schools fare under a high stakes testing regime? *Journal of Research in Rural Education*, 20(14), 1-12.
- Bond, L.A., & Cohen, D.A. (1991). The early impact of Indiana statewide testing for educational progress on local education agencies: Administrators' perceptions. In R.E. Stake (Ed.), *Advances in program evaluation* (pp. 75-99). Greenwich, CT: Jai.
- Borg, M.O., Plumlee, J.P., & Stranahan, H.A. (2007). Plenty of children left behind: High-stakes testing and graduation rates in Duval County, Florida. *Educational Policy*, 21, 695-716
- Brennan, R.T., Kim, J., Wenz-Gross, M., & Siperstein, G.N. (2001). The relative equitability of high-stakes testing versus teacher-assigned grades: An analysis of the Massachusetts comprehensive assessment system (MCAS). *Harvard Educational Review*, 71(2), 173-216.
- Brown, D.F. (1993). *The political influence of state testing reform through the eyes of principals and teachers*. (ERIC Document Reproduction Service No. Ed 360737)
- Bruner, J.S. (1966). After John Dewey, what? In R.D. Archambault (Ed.), *Dewey on education* (pp. 211-227). New York: Random House.

- Bryant, C. (2007). Planning and moderating focus group research. In M.V. Angrosino (Ed.), *Doing cultural anthropology* (pp. 115-127). Long Grove, IL: Waveland.
- Carnoy, M., & Loeb, S. (2002). Does external accountability affect student outcomes? A cross-state analysis. *Educational Evaluation and Policy Analysis*, 24(4), 305-331.
- Causey-Bush, T. (2005). Keep your eye on Texas and California: A look at testing, school reform, No Child Left Behind, and implications for students of color. *The Journal of Negro Education*, 74(4), 332-343.
- Cimbricz, S. (2002). *State-mandated testing and teachers' beliefs and practice*. Retrieved February 25, 2008 from <http://epaa.asu.edu/epaa/v10n18/>
- Costigan, A.T. (2002). Teaching the culture of high stakes testing: Listening to new teachers. *Action in Teacher Education*, 23(4), 28-34.
- Creswell, J.W. (1998). *Qualitative inquiry and research design: Choosing among five traditions*. Thousand Oaks, CA: Sage.
- Crocco, M.S., & Costigan, A.T. (2007). The narrowing of curriculum and pedagogy in the age of accountability. *Urban Education*, 42(6), 512-535.
- Cunningham, W.G., & Sanzo, T.D. (2002). Is high-stakes testing harming lower socioeconomic status schools? *NASSP Bulletin*, 86(631), 62-75.
- Curren, R. (2006). Connected learning and the foundations of psychometrics: A rejoinder. *Journal of Philosophy of Education*, 40(1), 17-29.
- Dantley, M.E. (2003). Purpose-driven leadership: The spiritual imperative to guiding schools beyond high-stakes testing and minimum proficiency. *Education and Urban Society*, 35(3), 273-291.
- Davis, A. (2006). High stakes testing and the structure of the mind: A reply to Randall Curren. *Journal of Philosophy of Education*, 40(1), 1-16.
- Dewey, J. (1915). *The school and society*. Chicago: University of Chicago Press.
- Dewey, J. (1916). *Democracy and education*. Middlesex, England: Echo Library.
- Dewey, J. (1938). *Experience and education*. New York: Macmillan.
- Domenech, D.A. (2000). My stakes well done. *The School Administrator*, 57(11), 16-19.
- Dorn, S. (2007). *Accountability Frankenstein: Understanding and taming the monster*. Charlotte, NC: Information Age.

- Escamilla, K., Chávez, L., & Vigil, P. (2005). Rethinking the “gap”: High-stakes testing and Spanish-speaking students in Colorado. *Journal of Teacher Education*, 56(132), 132-144.
- Escamilla, K., Mahon, E., Riley-Bernal, H., & Rutledge, D. (2003). High-stakes testing, Latinos, and English language learners: Lessons from Colorado. *Bilingual Research Journal*, 27(1), 25-49.
- Eisner, E.W. (1982). *Cognition and curriculum: A basis for deciding what to teach*. New York: Longman.
- Eisner, E.W. (1994). *The educational imagination: On the design and evaluation of school programs*. Upper Saddle River, NJ: Prentice Hall.
- Eisner, E.W. (1998a). *The enlightened eye: Qualitative inquiry and the enhancement of educational practice*. Upper Saddle River, NJ: Prentice Hall.
- Eisner, E.W. (1998b). *The kind of schools we need: Personal essays*. Portsmouth, NH: Heinemann.
- Eisner, E.W. (2002). *The arts and the creation of mind*. New Haven, CT: Yale University Press.
- Firestone, W.A., Mayrowetz, D., & Fairman, J. (1998). Performance-based assessment and instructional change: The effects of testing in Maine and Maryland. *Educational Evaluation and Policy Analysis*, 20(2), 95-113.
- Ganesh, T. G. (2002). Held hostage by high-stakes tests: Drawing as symbolic resistance. *Teacher Education Quarterly*, 29(4), 69-72.
- Gipps, C.V. (1994). *Beyond testing: Towards a theory of educational assessment*. Washington, DC: Falmer.
- Glasnapp, D.R., Poggio, J.P., & Miller, M.D. (1991). Impact of a “low stakes” state minimum competency testing program on policy, attitudes, and achievement. In R.E. Stake (Ed.), *Advances in program evaluation* (pp. 101-140). Greenwich, CT: Jai.
- Goebert, B., & Rosenthal, H. M. (2002). *Beyond listening: Learning the secret language of focus groups*. New York: Wiley.
- Goertz, M., & Duffy, M. (2003). Mapping the landscape of high-stakes testing and accountability programs. *Theory into Practice*, 42(1), 4-11.
- Goldin, C., & Katz, L.F. (2008). *The race between education and technology*. Cambridge, MA: The Belknap Press of Harvard University Press.

- Grant, S.G. (1996) Locating authority over content and pedagogy: Cross-current influences on teachers' thinking and practice. *Theory and Research in Social Education*, 24(3), 237-272.
- Grant, S.G. (2000). *Teachers and tests: Exploring teachers' perceptions of changes in the New York state testing program*. Retrieved June 22, 2007 from <http://epaa.asu.edu/epaa/v8n14.html>
- Grant, S.G. (2001). An uncertain lever: Exploring the influence of state-level testing in New York State on teaching social studies. *Teachers College Record*, 103(3), 398-426.
- Greenbaum, T.L. (1993). *The handbook for focus group research*. New York: Lexington.
- Greene, J.P., Winters, M.A., & Forster, G. (2003). Testing high stakes tests: Can we believe the results of accountability tests? *Civic Report*, 33, 1-17.
- Gunzenhauser, M.G. (2006). Normalizing the educated subject: A Foucaultian analysis of high-stakes accountability. *Educational Studies*, 39(3), 241-259.
- Halcomb, E.J., Gholizadeh, L., DiGiacomo, M., Phillips, J., & Davidson, P.M. (2007). Literature review: considerations in undertaking focus group research with culturally and linguistically diverse groups. *Journal of Clinical Nursing*, 16, 1000-1011.
- Hanson, F.A. (1993). *Testing testing: Social consequences of the examined life*. Los Angeles: University of California Press.
- Herman, J.L., & Golan, S. (1993). The effects of standardized testing on teaching and schools. *Educational Measurement: Issues and Practice*, 12(4), 20-25, 41-42.
- Hursh, D. (2005). The growth of high-stakes testing in the USA: Accountability, markets, and the decline in educational equality. *British Educational Research Journal*, 31(5), 605-622.
- Jayanthi, M., & Nelson, J.S. (2002). *Savvy decision making: An administrator's guide to using focus groups in schools*. Thousand Oaks, CA: Corwin.
- Johnson, D.D., & Johnson, B. (2006). *High stakes: Poverty, testing, and failure in American schools*. New York: Rowman & Littlefield.
- Jones, B.D., & Egley, R.J. (2006). Looking through different lenses: Teachers' and administrators' views of accountability. *Phi Delta Kappan*, 87(10), 767-771.
- Kamin, L. (1974). *The science and politics of IQ*. Potomac, MD: Erlbaum.

- Kleiber, P.B. (2004). Focus groups: More than a method of qualitative inquiry. In deMarras, K., & Lapan, S.D. (Eds.), *Foundations of research: Methods of inquiry in education and the social sciences* (pp. 87-102). Mahwan, NJ: Erlbaum.
- Koretz, D. (2008). *Measuring up: What educational testing really tells us*. Cambridge, MA: Harvard University Press.
- LaRoche, M.J., & Shriberg, D. (2004). High stakes exams and Latino students: Toward a culturally sensitive education for Latino children in the United States. *Journal of Educational and Psychological Consultation, 15*(2), 205-223.
- Lattimore, R. (2005). African American students' perceptions of their preparation for a high-stakes mathematics test. *Negro Educational Review, 56*(2/3), 135-146.
- Lee, J. (2006). Input-guarantee versus performance-guarantee approaches to school accountability: Cross-state comparisons of policies, resources, and outcomes. *Peabody Journal of Education, 81*(4), 43-64.
- Lloyd, G.M. (2007). Strategic compromise: A student teacher's design of kindergarten mathematics instruction in a high-stakes testing climate. *Journal of Teacher Education, 58*(4), 328-347.
- Mahon, E. (2006). High-stakes testing and English language learners: Questions of validity. *Bilingual Research Journal, 30*(2), 479-497.
- Manna, P. (2007). *School's in: Federalism and the national education agenda*. Washington, DC: Georgetown University Press.
- Massey, D. (2006). "You teach for me; I've had it!" A first-year teacher's cry for help. *Action in Teacher Education, 28*(3), 73-85.
- McDonnell, L.M. (2004). *Politics, persuasion, and educational testing*. Cambridge, MA: Harvard University Press.
- McNeil, L.M. (2000). *Contradictions of school reform: Educational costs of standardized testing*. New York: Routledge.
- Mehrens, W.A. (1998). *Consequences of assessment: What is the evidence?* Retrieved January 10, 2008 from <http://epaa.asu.edu/epaa/v6n13.html>.
- Miles, M.B., & Huberman, A.M. (1994). *Qualitative data analysis*. Thousand Oaks, CA: Sage.
- Mitchell, R. (1992). *Testing for learning*. New York: The Free Press.
- Morgan, D.L. (1997). *Focus groups as qualitative research*. Newbury Park, CA: Sage.



- Morrison, D. E. (1998). *The search for a method: Focus groups and the development of mass communication research*. Luton, Bedfordshire, UK: University of Luton Press.
- National Commission on Excellence in Education. (1983). *A Nation at risk*. Washington, DC: U.S. Department of Education.
- Neill, M. (2003). High stakes, high risk. *American School Board Journal*, 190, 18-21.
- Paris, S.G., & Urdan, T. (1994). Teachers' perceptions of standardized achievement tests. *Educational Policy*, 8(2), 137-156.
- Paris, S.G., & Urdan, T. (2000). Policies and practices of high-stakes testing that influence teachers and schools. *Issues in Education*, 6(1/2), 83-108.
- Popham, W.J. (1999). Where large scale educational assessment is heading and why it shouldn't. *Educational Measurement: Issues and Practice*, 18(3), 13-17.
- Popham, W.J. (2000). The mismeasurement of educational quality. *The School Administrator*, 57(11), 12-15.
- Pringle, R.M., & Martin, S.C. (2005). The potential impacts of upcoming high-stakes testing on the teaching of science in elementary classrooms. *Research in Science Education*, 35, 347-361.
- Raphael, T.E., & Au, K.H. (2005). QAR: Enhancing comprehension and test taking across grades and content areas. *The Reading Teacher*, 59(3), 206-221.
- Raymond, M.E., & Hanushek, E.A. (2003). High-stakes research. *Education Next*, 3(3), 48-55.
- Reis, H.T., & Wheeler, L. (1991). Studying social interaction using the Rochester interaction record. In M.P. Zanna (Ed.), *Advances in experimental social psychology* (pp. 269-318). San Diego, CA: Academic.
- Rothman, R. (1995). *Measuring up: Standards, assessment, and school reform*. San Francisco: Jossey-Bass.
- Sacks, P. (1999). *Standardized minds*. New York: Da Capo.
- Sacks, P. (2000). Predictable losers in testing schemes. *The School Administrator*, 57(11), 6-9.

- Seidman, I. (2006). *Interviewing as qualitative research: A guide for researchers in education and the social sciences*. New York: Teachers College Press.
- Sirotnik, K.A. (2002). Promoting responsible accountability in schools and education. *Phi Delta Kappan*, 83(9), 662-673.
- Sloan, K. (2007). High-stakes accountability, minority youth, and ethnography: Assessing the multiple effects. *Anthropology and Education Quarterly*, 38(1), 24-41.
- Smith, E. (2005). Raising standards in American schools: the case of No Child Left Behind. *Journal of Education Policy*, 20(4), 507-524.
- Smith, M.L. (1991). Put to the test: The effects of external testing on teachers. *Educational Researcher*, 20(5), 8-11.
- Smith, M.L. (2000). Validity and accountability in high-stakes testing. *Journal of Teacher Education*, 51(5), 334-344.
- Smith, M.L. (2004). *Political spectacle and the fate of American schools*. New York: Routledge Falmer.
- Stake, R.E., & Theobald, P. (1991). Teachers' views of testing's impact on classrooms. In R.E. Stake (Ed.), *Advances in program evaluation* (pp. 189-201). Greenwich, CT: JAI.
- Stoskopf, A. (2000). The forgotten history of eugenics. In Swope, K., & Miner, B. (Eds.). *Failing our kids: Why the testing craze won't fix our schools* (pp. 76-80). Milwaukee, WI: Rethinking Schools, LTD.
- Sullivan, R.G. (1991). Teachers' perceptions of the effects of testing on classroom practices. In R.E. Stake (Ed.), *Advances in program evaluation* (pp. 145-161). Greenwich, CT: JAI.
- Tittle, C.K., Kelly-Benjamin, K., & Sacks, J. (1991). The construction of validity and effects of large-scale assessments in the schools. In R.E. Stake (Ed.), *Advances in program evaluation* (pp. 223-254). Greenwich, CT: JAI.
- Triplett, C.F., & Barksdale, M.A. (2005). Third through sixth graders' perceptions of high-stakes testing. *Journal of Literacy Research*, 37(2), 237-260.
- Tyack, D., & Hansot, E. (1982). *Managers of virtue: Public school leadership in America, 1820-1980*. New York: Basic.
- Valenzuela, A. (Ed.) (2005). *Leaving children behind: How "Texas-style" accountability fails Latino youth*. Albany: State University of New York Press.

- Valli, L., & Buese, D. (2007). The changing roles of teachers in an era of high-stakes accountability. *American Educational Research Journal*, 44(3), 519-558.
- Van Hover, S., Hicks, D., & Irwin, W. (2007). Beginning teachers thinking historically? Negotiating the context of Virginia's high-stakes tests. *International Journal of Social Education*, 22(1), 85-114.
- Watanabe, M. (2007). Displaced teacher and state priorities in a high-stakes accountability context. *Educational Policy*, 21(2), 311-368.
- Weaver, P.E. (2004). The culture of teaching and mentoring for compliance. *Childhood Education*, 80(5), 258-260.
- Wiley, T.G., & Wright, W.E. (2004). Against the undertow: Language-minority education policy and politics in the "Age of Accountability." *Educational Policy*, 18(1), 142-168.
- Willis, D. (1993). Learning and assessment: Exposing the inconsistencies of theory and practice. *Oxford Review of Education*, 19(3), 383-403.
- Wilson, B.L., & Corbett, L. (1991). Two state minimum competency testing programs and their effects on curriculum and instruction. In R.E. Stake (Ed.), *Advances in program evaluation* (pp. 7-40). Greenwich, CT: JAI.
- Wolcott, H.F. (1994). *Transforming qualitative data: Description, analysis, and interpretation*. Thousand Oaks, CA: Sage.
- Wright, W.E. (2005). English language learners left behind in Arizona: The nullification of accommodations in the intersection of federal and state policies. *Bilingual Research Journal*, 29(1), 1-29.
- Zancanella, D. (1992). The influence of state-mandated testing on teachers of literature. *Educational Evaluation and Policy Analysis*, 14(3), 283-295.

## VITA

## JOSEPH L. HEYWOOD

(435) 590-8924  
joseph.heywood@gmail.com

455 Carlos Ct  
Moab UT 84532

---

**Education**

Ed.D. in Curriculum and Instruction  
Utah State University – December 2009

M.Ed. in Educational Leadership  
Northern Arizona University – June 2002

BA in History Teaching with English Teaching minor  
Brigham Young University – August 1999

**Employment**

Director/Principal  
Moab Charter School, Moab UT; July 2008 to present

Dean of Instruction  
Window Rock High School, Fort Defiance AZ; July 2007 to June 2008

High School English Teacher  
Juab High School, Nephi UT; August 2003 to June 2007

Junior High School English Teacher  
Hesperia Junior High, Hesperia CA; Sept 2002 to June 2003

Adjunct Instructor of English, Reading, and History  
Diné College, Ganado AZ; August 2000 to May 2002; August 2007 to December 2007

High School English and History Teacher; History Department Chair  
Ganado High School, Ganado AZ; August 2000 to May 2002

Middle School History and English Teacher  
Southridge Middle School, Fontana CA; Sept 1999 to August 2000

## **Education-Related Duties**

### ***Director/Principal, Moab Charter School***

Responsibility for all school programs, staffing, evaluations, school budget, testing, discipline, and the special education program. Charged with carrying out the directives of the governing board and establishing and maintaining cooperative relationships with all Utah State Office of Education departments and programs.

### ***Dean of Instruction, Window Rock High School***

Responsibility for English Language Learner program, all state and district testing, teacher evaluations, all tutoring programs, summer school development, textbook and curriculum material ordering, as well as assisting with parent meetings, discipline issues, and other general school administration issues.

### ***Snow College Concurrent Enrollment Instructor, Juab High School***

Taught an English 1010 class for college credit through a local community college. The class was taught at the high school.

### ***District Curriculum Blocking Project Participant, Juab School District***

Served as a two-year participant on a team which blocked the entire curriculum for the high school English Department. In addition, I designed the end-of-quarter comprehensive exams for all sophomore English classes. I also formatted the tests using computer software for internet delivery and analysis of the exams for all sophomore and junior English courses.

### ***School Newspaper Advisor, Juab High School***

Taught the journalism class and supervised the school newspaper. We published a monthly edition of 12 pages with contributions from over 30 student staff members.

### ***Literary Magazine Advisor, Juab High School***

For three years, I served as the literary magazine advisor. Along with a small student staff, we printed, bound, and sold student essays, poems, and short stories on an annual basis.

### ***Adult Evening School Instructor, Juab High School***

I taught an adult English class for students working toward passing the G.E.D. or seeking high school credit. I designed and delivered the course.

### ***After-School Clinic Instructor, Hesperia Junior High***

For one year, I taught standardized remedial courses to eighth grade students who had failed to meet benchmark requirements on English exams.

### ***Designed courses and taught at Diné College***

Over a two-year period, I designed and taught 8 courses in English, Southwestern Literature, World History, and Reading for Native American college students and working adults on the Navajo Reservation in Arizona. In addition, I designed and taught courses the Fall of 2007 in reading and composition.

### ***History Department Chair, Ganado High School***

At the request of my colleagues, I served as the History Department Chair, leading our department through an important textbook and curriculum acquisition year.

***School Site Council Member, Ganado High School***

For one year, I served as the School Site Council member representing the History Department at Ganado High. This administrative team met regularly to discuss all major school decisions with votes and recommendations governing administrator and department action.

***At Risk-Retention Intervention Instructor, Southridge Middle School***

For one year, I taught remedial courses for struggling English students. The courses were semi-structured with much of the lesson planning left to the teachers.

***Alternative to Suspension Coordinator, Southridge Middle School***

For one year, I served as the Alternative to Suspension Coordinator for after-school detention. These students had broken school rules and were required to serve the school with a certain number of after-school hours rather than missing school time due to suspension.

**Professional Presentations**

Utah Coalition for Educational Technology Conference, March 2007, Presenter  
“Using PowerPoints as Assessments”