Teacher insights on high-stakes standardized assessments: the impact of school reform policy on the classroom

Jennifer Timmer
Louisiana State University and Agricultural and Mechanical College
TEACHER MOTIVATION AND HIGH-STAKES STANDARDIZED ASSESSMENTS: THE IMPACT OF SCHOOL REFORM POLICY ON THE CLASSROOM

A Thesis

Submitted to the Graduate Faculty of the Louisiana State University and Agricultural and Mechanical College in partial fulfillment of the requirements for the degree of Master of Arts in The Department of Educational Theory, Policy, and Practice

by

Jennifer Dee Timmer
B.A., University of Kansas, 2002
M.A., University of Miami, 2005
May 2012
Acknowledgements

I am grateful to all who have assisted me in this endeavor. In particular, I want to recognize and thank Petra Munro Hendry, Keena Arbuthnot, Jacqueline Bach, Tracy Rumans, John Timmer, Barbara Timmer, and David Schenck. Many thanks for the support, encouragement, thoughtful insights, and lively discussion along the way.
Table of Contents

ACKNOWLEDGEMENTS ........................................................................................................ ii

LIST OF TABLES .................................................................................................................. iv

ABSTRACT .............................................................................................................................. v

CHAPTER 1: INTRODUCTION ............................................................................................... 1
   INTRODUCTION .................................................................................................................. 1
   PURPOSE AND RESEARCH QUESTIONS ........................................................................... 6

CHAPTER 2: REVIEW OF LITERATURE ................................................................................. 9
   HISTORICAL OVERVIEW OF ASSESSMENT IN EDUCATION ........................................... 9
   CONTEMPORARY RELATED STUDIES ............................................................................. 38
   CONCLUSION .................................................................................................................... 53

CHAPTER 3: METHODS ......................................................................................................... 55
   PROCEDURES ................................................................................................................... 55
   PARTICIPANTS ................................................................................................................ 56
   INSTRUMENTS ............................................................................................................... 58
   CONSTRUCTS ................................................................................................................ 59
   DATA ANALYSIS .......................................................................................................... 61

CHAPTER 4: RESULTS ........................................................................................................... 62
   OVERALL FINDINGS ....................................................................................................... 62
   GROUP COMPARISONS ................................................................................................. 63

CHAPTER 5: CONCLUSION .................................................................................................... 68
   OVERALL RESULTS ....................................................................................................... 68
   DISCUSSION .................................................................................................................. 74
   LIMITATIONS .............................................................................................................. 77

REFERENCES ...................................................................................................................... 79

APPENDIX A: SURVEY ......................................................................................................... 83

VITA ....................................................................................................................................... 89
List of Tables

1. Participant Demographics .................................................................57
2. Sample Subgroup Demographics ..........................................................58
3. Gender Comparison .............................................................................63
4. Training Comparison ............................................................................64
5. Experience Comparison .......................................................................65
6. Gifted and Talented Student Work Comparison ...................................67
Abstract

Education reform policies have focused on high-stakes assessments primarily utilizing standardized tests for accountability purposes. As testing practices have evolved due to a variety of factors throughout the past century or so, they have become fully integrated into public schooling in the United States. These tests are having a marked impact on teachers and education, as teachers feel pressure to produce exceptional student results and modify their instructional practices, often “teaching to the test” and narrowing the curriculum in order to focus on the requirements of the mandates. This study examines a survey of public school teachers to ascertain their experiences and perspectives regarding the impact of these testing policies. The results show that teachers are feeling an immense amount of pressure, their instructional planning and classroom practices are impacted by the tests, and they do not find the tests particularly helpful. Differences between teachers’ gender, education, experience, subject focus, and work with gifted and talented students are also examined. As testing is likely to be a force in public education in America for the foreseeable future, the results of this study can be used to ascertain teacher concerns and develop supports for teachers to help address those concerns.
Chapter 1
Introduction

School curricula have always been evolving as societal, cultural, and political shifts impact a community’s sense of learning and education. Recently, education has become strongly linked with assessments and student achievement, and, more specifically, standardized tests, providing numerical data ostensibly measuring the learning of each student and easily compared to other students. From *A Nation at Risk* (1983) to *No Child Left Behind* (2002) and *Race to the Top* (2012), the discourse surrounding education in America has centered on education as a competition critical to global economic success complete with winners and losers. Testing is at the heart of the reform efforts geared towards solving the problem of substandard performance. While states have utilized assessments for evaluative and even high-stakes purposes for many years, over the past decade, through national mandates, poor standardized test performance has become more closely linked with high-stakes accountability consequences, with students facing the threat of not being promoted to the next grade, teachers worrying about losing their jobs or opportunities for compensation, and schools confronting the possibility of being taken over or shut down entirely.

With student and school performance being monitored so closely and with such important implications attached, teachers have found themselves under intense scrutiny. Teachers must balance the pressure to improve student performance on these tests with their own professional judgment of their students’ educational needs. As teachers are frequently noted to be the most
important in-school factor in student performance,\textsuperscript{1} it is critical to understand their experiences of merging the requirements of the testing mandates with student instruction. While policy makers, school boards, and administrators determine the curricular directives, teachers and students are the individuals who actually implement the policies and participate in the lived experience of learning within the classroom. Understanding teachers’ perspectives of testing and how it impacts their teaching and classrooms helps illustrate the true consequences of the testing movement. Further, understanding this impact can assist educators at all levels in developing and implementing effective instructional practices.

In appreciating the history of testing and evaluation in American schooling, we can understand the context of our current assessment practices and examine both the strengths and weaknesses of current evaluative structures. A variety of factors, educational, scientific, cultural, and military, for example, have all impacted the current role of testing and assessment as practiced in public schools in the United States. The notion that we should be able to effectively evaluate our students, as well as their teachers and schools, makes sense, as the importance of educational success cannot be denied, and our culture has embraced the notion of verifiable proof in understanding the success or failure of social institutions. However, do high-stakes standardized tests truly provide this evidence of learning we seek? The answer to that question is complicated and difficult to answer. Standardized tests can provide some data regarding student learning and achievement, but they are only one form of assessment. The tests repeatedly demonstrate an achievement gap between students of differing socioeconomic statuses and races, with students from high poverty schools and students of color earning lower scores than their

\textsuperscript{1} For an overview of the research surrounding the notion that teachers are critical in-school factors of student performance, see Linda Darling Hammond’s \textit{The Flat World and Education: How America’s Commitment to Equity Will Determine Our Future} (2010), p. 43.
more affluent counterparts, often attending schools with low-minority enrollment (Arbuthnot, 2011; Darling-Hammond, 2010; Meier & Wood, 2004; Moss, Pullin, Gee, Haertel, & Young, 2008; Ravitch, 2011; Taubman, 2009; Winfield, 2007). Questions of test validity and appropriateness for use in high-stakes contexts are pressing and too frequently overlooked. These concerns add to the challenges teachers face in working with these policies.

In order to understand the role of testing in contemporary schools, it is important to examine the history of assessing and measuring students and the intersections of testing and education policy. Teachers have essentially always been assessing students, at least informally, as they deliver instruction and ascertain students’ educational needs. However, the development and implementation of formal, standardized assessments has rapidly accelerated and overtaken education in America in the form of data-driven assessment and numerical quantification of student learning. As the profession of teaching became more feminized, teachers were seen as needing male supervision, those with decision-making power, and the locus of control shifted outside the classroom (Lagemann, 2000). At the same time, a growing interest was developing in the advances in science and technology, and public education in America fully engaged in the testing boom in education throughout the twentieth century, a measure of assessment created outside the classroom and out of the teacher’s control.²

Hopes for societal improvement and success in global competition have often been the impetus for more testing and assessment. Assessment has now both broadened in usage and narrowed in format, in particular, the standardized assessments most often discussed in contemporary discourse. Policy mandates now require high-stakes standardized assessments in

all public schools in America and attach accountability measures to student performance. Over the course of the twentieth century, assessment became a primary feature of schooling in America, and is the cornerstone of the school reform initiatives dominating educational discourse today (Darling-Hammond, 2010; Pinar, Reynolds, Slattery, & Taubman, 1995; Ravitch, 2010b; Taubman, 2009). Based on these accountability standards, schools are labeled as failing, only adding to the negativity of the discourse surrounding education. Various reform efforts are often targeted toward the privatization of education, the neoliberal response often dominating the political discourse.

Teachers and students, the people engaging in the lived experience of education in the classroom, face the impact of testing on schooling most intensely. Research on the impact of testing on teachers and schooling has typically focused on the prevalence of “teaching to the test” and a narrowed curriculum, as teachers devote time to tested subjects and test preparation skills (Darling-Hammond, 2010; Kocabas, 2009; Martin, 2006; Ravitch, 2010b; Sheldon & Biddle, 1998; Skaalvik & Skaalvik, 2011; Urhahne, Chao, Florineth, Luttenberger, & Paechter, 2011). Surveys of teachers are frequently used to ascertain their perspectives on their experiences with testing and can show what particular concerns are most pressing. They can highlight what areas of education teachers feel are most impacted by the standards and accountability movement, and how. As an examination of the history and development of testing seems to indicate, testing as a tool for assessment is likely to continue to have some influence for the foreseeable future. There is then great value in identifying the consequences of the mandates, both problematic and beneficial, and then using that information to help teachers.

Much of the research conducted regarding testing and its impact on schooling has focused on several broad areas: testing and test validity (Arbuthnot, 2011; Taubman, 2009),
teachers’ opinions of testing (Buck, Ritter, Jensen, & Rose, 2010; Craig, 2010; James & Pedder, 2006; Monsaas & Engelhard, 1993), the pressure teachers feel to improve student performance (Amrein-Beardsley, Berliner, & Rideau, 2010; Moon, Brighton, Jarvis, & Hall, 2007; Mulvenon, Stegman, & Ritter, 2005; Perreault, 2000), the impact of the tests on instructional time and planning (Darling-Hammond, 2010; Faulkner & Cook, 2006; Musoleno & White, 2010; Wills & Sandholtz, 2009; Zellmer, Frontier, & Pheifer, 2006), and the support school administrative personnel provide in the testing process (DeMoss, 2002; Reitzug, West, & Angel, 2008; Singh & Al-Fadhli, 2011; Sunderman, Orfield, & Kim, 2006). Teacher surveys are frequently used, as are qualitative studies involving observation and/or interviews.

Overwhelmingly, these studies have found that teachers are not opposed to student assessment in theory, as they believe it is critical to informing instructional practice in their classrooms. However, they tend to find that the implementation of high-stakes evaluative policies actually detracts from the learning environment. Teachers tend to report feeling high levels of pressure. Further, due to this pressure, teachers are open in reporting a significant amount of instructional time geared towards teaching their students how to perform well on the tests, even when these practices do not align with their own educational values. Many teachers even admit to having participated in unethical practices in order to achieve the high test scores required by their schools and state and federal mandates (Amrein-Beardsley et al., 2010). Collaborative efforts between teachers and administrators seem to help the teachers, but are not always present in the schools, leading back to the problematic situations in which teachers find themselves due to high-stakes testing policies.
Purpose and Research Questions

The policies of the No Child Left Behind act have now been in effect for just over a decade, and a new understanding of the impact of their implementation can be found. The purpose of this study is to investigate teachers’ opinions of the impact of the mandated high-stakes standardized assessments on their schools and teaching. By surveying public school teachers, we can learn how those inside the classroom are experiencing the policies they must implement.

Through investigating these teacher perspectives, this study seeks to understand the impact of high-stakes standardized assessments on teachers’ classroom practices. The following research questions will be examined:

- What are teachers’ opinions regarding the helpfulness of high-stakes standardized assessments in their teaching, and how do they differ based on teachers’ gender, levels of education, experience, and work with gifted and talented students?
- What are the effects of high-stakes standardized assessments on teachers and their instruction, and are there differences based on teachers’ gender, levels of education, experience, and work with gifted and talented students?
- To what extent do teachers feel pressure to ensure their students perform well on these tests, and does that differ based on teachers’ gender, levels of education, experience, and work with gifted and talented students?
- Do teachers feel supported by their school administration in their endeavors to help students succeed in the classroom and on these assessments, and does their sense of support differ based on teachers’ gender, levels of education, experience, and work with gifted and talented students?
Through surveying public school teachers at a school working to improve its standing on the state-mandated assessments, this study hopes to learn about the experiences of teachers regarding the influence they feel standardized assessments have on their teaching, classrooms, and school. The mandates of NCLB have now been implemented for a full decade and the deadline for 100% proficiency in 2014 is fast approaching. While a fair amount of research has investigated these teacher responses, programs and initiatives intended to improve student performance are now theoretically more integrated into school curricula. Teacher responses are not anticipatory or presumptive, but are based on their daily experiences in the classroom and in this high-stakes testing culture society has defined. Thus, learning about teachers and their experiences in the present day is especially relevant. Teacher insights are valuable, as their experiences reflect the real impact of policy implementation.

The results of this study can be beneficial for policy makers at all levels, district and school administrators, and teachers. Policy makers must seek a deep understanding of the impact of high-stakes standardized assessments on teaching and learning in order to develop and implement policies that positively impact schools and truly measure student achievement and meaningful learning. Creating policies from a distance requires a thorough investigation of the impact of those policies on the people involved in their implementation on a daily basis.

Administrators are the primary support for teachers as policies are implemented at the school level; therefore, administrators will also benefit from learning about the perspectives of the teachers they support. This will assist them in creating training, support, and evaluation programs for teachers that will genuinely be helpful. Finally, teachers’ own understandings of their experiences with high-stakes standardized assessments can help inform their instructional
planning and classroom time, as well as assist them in contextualizing their experiences and expectations.


Chapter 2
Review of Literature

Historical Overview of Assessment in Education

Contemporary discourse regarding public education in America focuses on failing schools that leave our students ill-prepared to enter the global marketplace and maintain America’s standing in the world as an economic power. An increasingly loud voice within the discourse calls for the abolishment of public education altogether, especially any federal presence. However, many Americans still believe that public schooling is necessary for an equitable and democratic society, and that some federal regulation is necessary to ensure students in all states receive appropriate schooling, and so “our faith in public schools as the great equalizer remains” (Meier & Wood, 2004, p. viii). Thus, debate regarding school reform and the policies impacting it rages on. The policies enacted, whether outlining involvement or abandonment by the federal government, will no doubt have a drastic impact on education and teaching in America, thus demonstrating the importance of understanding the history of policy, as well as the various factors at play in defining key decisions.

Education in America is not a stagnant field but instead has been evolving as a variety of social, cultural, political, and economic factors have impacted American society and its educational wants and needs. Appreciating this expansive history allows for a more thoughtful examination of contemporary education concerns, as viewing curriculum and learning in an ahistorical manner leads to a sense of urgency and crisis, rather than as an ever-evolving and developing field that will change with societal shifts (Pinar et al., 1995). Current educational policy emphasizes testing and assessment, placing high-stakes consequences for schools and teachers based on student performance, which is the result of decades of influence from a variety
of fields, such as psychology, economics, and the military, among others. This reliance on testing informs our thought and discourse. As America now administers over 100 million standardized assessments each year (Taubman, 2009), a number only growing, testing overwhelms educational discourse and is more powerful than ever (Moss et al., 2008). But where do we get the notion that testing is the primary indicator of learning, and that it should define education and schooling? The reliance on testing implies that education is a science, a notion that has been debated by many educators, scientists, and researchers (Lagemann, 2000). And whether or not education is a science, how can one determine whether or not students are learning, and how precise can those determinations be? Can all learning be measured? Testing and assessment undoubtedly play a role in successful education, but what that role should be has been, and will continue to be, debated. The question and debate are of utmost importance, as the answers we find to those questions determine the future of teaching and education in America and, in a sense, the future of America itself.

**Early Theories of Learning and Science.**

As public schooling in America was beginning to develop and spread throughout the nineteenth century, a variety of influences and beliefs can be identified that worked to shape education. Learning was thought to be demonstrated by memorization and repetition; the mind was thought to be a muscle that would develop from repeated exertion, much like any other muscle in the body that might develop from repeated exercise (Pinar et al., 1995). Also important early was the influence of humanism, which led to standardization, as the focus was on the person in the abstract, rather than as an individual (Pinar et al., 1995). This led to the idea
that a particular “method” would advance learning, a generalized curriculum which could be applied to all children.

In the late nineteenth century, G. Stanley Hall extended the notion of the abstract learner in developing his stage theory of child development, in which children experience a progression of developmental stages as they grow. His work dramatically shaped schooling to this day, as it identified age groupings for appropriate school structuring. However, Hall also embraced the child study movement in informing his work and focused on individualization as well. He focused on the scientific study of the child and argued that heredity, not environment, was the biggest factor in understanding the child.

The study of eugenics was also gaining traction during this time in the wake of Francis Galton and other scientists studying the possibilities of scientific, biological engineering (Gillham N., 2009; Lorimer, 1990; Sandall, 2008; Winfield, 2007). Eugenics was seen as a possible method of improving society. Galton’s work led to many advances in statistics and established the fields of educational psychology and psychometrics (Sandall, 2008). The notion of testing for intelligence has since spread and become commonplace, although IQ tests’ roots in eugenics are often either forgotten or ignored. The development of IQ tests spawned the creation of a massive influx of tests and measures throughout the the 1920s and 1930s, similar to the boom being experienced today (Lagemann, 2000; Taubman, 2009). Hall was in alignment with the view of education as involving measurement and testing (Lagemann, 2000; Pinar et al., 1995).
**Progressivism.**

The debate between the child study and scientific approaches has been a continuing divide in education, stemming from the scientific research and the Progressive movements in the early twentieth century. Most famously associated with John Dewey, Progressivism focused on the ultimate goal of education as a force for democracy and social change, teaching children how to enter and embrace the larger community surrounding them. Curriculum was thus integrated with children’s lives and experience as the focus within the classroom. Progressivism was child-centered, focused on the experiences of the student. The teacher was a facilitator of student learning, engaging in democratic classroom practices that allowed for exploration and discovery (Dewey, 1990; Pinar et al., 1995).

However, Progressivism was not divorced from science. Dewey himself saw education as improved through a laboratory of sorts, in which one could examine scientific work in a practical setting (Dewey, 1990). Further, Dewey believed that “educational inquiry should be directed toward finding ways to increase educational efficiency” (Lagemann, 2000, p. 49). Dewey also believed schools and teachers should be held accountable for student learning, although this accountability was community-based, as he believed students needed to learn how to engage in society. Schools should thus not operate in isolation from the community and teachers were responsible for enabling those interactions (Dewey, 1990). Progressivism maintained a sense of humanity and a focus on the person as participant in education and schooling.

The debate continues as to the impact Progressivism has had on education in America, with many claiming that science and the social efficiency movements overwhelmed Progressivism’s influence. However, its ideas continue to surface in curricular debate. While
the tenets of Progressivism continue to be discussed regarding curriculum, though, the notion that learning can and should be measured through regular testing has overshadowed all other curricular debate, leaving room for discussion regarding how to teach, but less so regarding what to teach. From the early stages of debate between the two factions, scientific study has fit more conveniently with other areas of societal interest, mainly science and technology, and how they can lead to social progress.

**Edward L. Thorndike and Experimental Psychology.**

A major contributor to the development of education as connected to psychology and science was Edward Thorndike, an experimental psychologist who wanted to utilize the research methods of the physical sciences in analyzing education, thus providing objectivity and verifiability (Pinar et al., 1995). He saw education as a form of “human engineering” with training at its core. Thorndike’s behaviorist views involved a set stimulus-response connection, with learning being the response to the appropriate stimulus. The behavior could then be observed, which easily transitioned to being measurable (Taubman, 2009). He thus believed that “human experience would then be mathematicized” (Pinar et al., 1995, p. 92). Not only does this emphasize learning as an abstract process that can be mechanized, it also underscores the foundations of understanding learning as something that can be understood through statistics and numerical representations of knowledge. The mind is thus a machine, rather than an individual human conception and understanding of the world. This posits a very different conception of the teacher from the Progressive notion. Here, the teacher is more of a trainer, developing appropriate behaviors, rather than a community-builder.
Thorndike’s analysis of students and learning hinted at the importance of efficiency, as ever more efficient modes of teaching could be used to achieve the learning demonstrated in student response. However, one can already see the potential for problems when students memorize the correct response in a given situation without understanding why that response is accurate. In this situation, has true learning occurred? Thorndike was not alone in seeking efficiency; the National Education Association formed a committee to study ways to increase efficiency in schools in 1911. This took place along with the development of the influential eugenics movement seeking scientific bases for categorizing and segregating people based on race. A growing acceptance of natural evolution led to a belief that people were predisposed either to be able to learn or not. Thorndike emphasized “the importance of basing educational studies on controlled experimentation and precise quantitative measurements” (Lagemann, 2000, p. 59), while also believing that, “What anyone becomes by education… depends on what he is by nature” (qtd. in Lagemann, 2000, p. 58). Thus, Thorndike’s work helped shape a discourse of education focused on innate ability, measurable by “reliable” scientific methods.

**Scientific Management, Taylor, and Bobbitt.**

Perhaps one of the most important shapers of education and testing was Frederick Taylor, who claimed that scientific management could guarantee the efficiency and effectiveness sought based on economic principles (Pinar et al., 1995). An important element of his approach was task analysis, which was the breakdown of a process into smaller defined pieces, which could then be analyzed for maximum efficiency, leading to a more efficient process overall. This led to the division of curriculum into segments that could be sequenced for enhanced instruction.
The goal of this curriculum breakdown was social utility, as students who went down the “assembly line” would become “socially useful citizens” (Pinar et al., 1995, p. 95).

Franklin Bobbitt embraced the importance of efficiency but also incorporated some ideals of Progressivism. He saw school as preparation for the adult world, much as the Progressives saw it as the space for educating children on how to participate in society and democracy. However, he also included concepts of task analysis in his application, finding that the work of educators and curriculum creators must be to identify specific tasks that will prepare children for this engagement (Pinar et al., 1995). Bobbitt’s work was in a way a hybrid, as he saw a complex society that required a more advanced education. He wanted children to engage in experiences that would train them to approach the tasks necessary to adulthood (Lagemann, 2000). Even as he exhibited some Progressive ideals, they are tempered by his need for scientific analysis and implementation in order to effectively serve society. These early educators and scientists thus constructed a debate and discourse laying the groundwork for the requirement of data in verifying the success of education, a requirement that has more recently been expanded to increase assessment and testing data.

**The Eight-Year Study.**

In the 1930s, the Progressive Education Association conducted a study that came to be known as the Eight-Year Study, which investigated secondary schools and their students to determine how schools served those students and how prepared those students were for the rigors of collegiate study. Schools volunteered to participate in the study and teachers and administrators revised their curricula in experimental ways to explore what might best prepare secondary students for participation in the opportunities ahead of them. The research questions
focused on the relevancy and significance of the high school curriculum to students’ lives, questions still debated today (Pinar et al., 1995; Winfield, 2007). Ralph W. Tyler was the Director of Research for the Evaluation Staff, and the importance of his participation in the study and in the development of curriculum afterwards cannot be ignored. Tyler was a strong proponent of the scientific study of education, a strong theme in the discourse of the time. Further, he emphasized evaluation and measurement, applying the earlier understanding of learning as that which is observably demonstrated on an assessment (Pinar et al., 1995). Importantly, though, educators were at the center of curricular design.

The results of the study demonstrated that there was no fixed pattern for success, however much one was anticipated and desired. While somewhat overlooked due to the national focus on the pending involvement of the United States in World War II, the Eight-Year Study found that “success in college is not dependent on credits earned in high school in the traditionally prescribed subjects” (Pinar et al., 1995, p. 137). Further, the most experimental schools were actually found to be the most successful, even though they were divorced from earlier concepts of required curriculum and method. Despite the negligible impact of the results on curricular discussion due to the preoccupation of the country in WWII, the sense that experimentation in education is valuable is still seen today, as contemporary discourse often discredits traditional schooling in favor of experimental charter schools and other alternatives. That schools might prove more effective given the space for experimentation, though, does not seem to translate into an appreciation for other creative and innovative school- and classroom-level approaches and interventions, even though the Eight-Year Study was based on individual schools’ freedom to design their own curriculum.
**Basic Principles of Curriculum and Instruction and the Tyler Rationale.**

Tyler was instrumental in defining curriculum and evaluation as it is known today. In his seminal work *Basic Principles of Curriculum and Instruction*, Tyler outlines his overall goal:

> Curriculum planning is a continuous process and... as materials and procedures are developed, they are tried out, their results appraised, their inadequacies identified, suggested improvements indicated; there is replanning, redevelopment and then reappraisal; and in this kind of continuing cycle, it is possible for the curriculum and instructional program to be continuously improved over the years. In this way we may hope to have an increasingly more effective educational program rather than depending so much upon hit and miss judgment as a basis for curriculum development. (Tyler, 1949, p. 123)

Teachers were thus responsible for continual curriculum development and revision based on student progress, placing them in an important position within education. Tyler believed this reappraisal and improvement would be facilitated by evaluation of student progress. His proposals were highly influential; “The simplicity and functionality of the Tyler Rationale were compelling for many educators” (Pinar et al., 1995, p. 149). Tyler’s focus was on outlining precise objectives for designing a curriculum which are then sequenced appropriately and evaluated. This requires well-defined subjects so that specific objectives can be achieved (Tyler, 1949). These precise objectives siphoned from the larger curricular categories resemble the tasks identified through task analysis, an earlier partitioning of knowledge for creation of measureable objectives. Tyler felt the objectives needed to be clearly defined “so as to provide a concrete guide in the selection and planning of learning experiences.... It is absolutely essential that they be defined in order to make an evaluation since unless there is some clear conception of the sort of behavior implied by the objectives, one has no way of telling what kind of behavior to look for in the students in order to see to what degree these objectives are being realized” (Tyler, 1949, p. 111). Thus, the outlining of specific learning objectives leads to a solid curricular plan,
which leads to an evaluation plan that assesses students’ ability to express their learning. While teachers here play an important role, some standardization of objectives and evaluation is also present.

Tyler’s understanding of evaluation was a basic question asking whether or not the instruction is producing results. Specifically, he writes that “evaluation then becomes a process for finding out how far the learning experiences as developed and organized are actually producing the desired results and the process of evaluation will involve identifying the strengths and weaknesses of the plans” (Tyler, 1949, p. 105). While his focus on evaluation has been extended to our current broad application of testing, he actually seems to be pointing more so to the importance of evaluating the instructional plan rather than the student. Tyler continually points to the importance of the teacher evaluating and revising the curriculum and instructional plan. If the student does not exhibit the desired outcome, Tyler does not focus on blaming any of the people involved, but rather the system and its implementation. Formative evaluation is central to this notion of education and curriculum. Further, while he favored a system of annual student evaluation, his description of this evaluation included a variety of student products and artifacts, such as observations, interviews, questionnaires, student products, and school records such as library records and cafeteria habits (Tyler, 1949). His sense of evaluation was thus much broader than often assumed.

Still, though, he felt traditional tests were also an appropriate method for measuring student learning. These tests, however, must be the proper instruments for measurement, and instruments should only be selected after the learning objectives have been identified (Tyler, 1949). This is a distinct contrast to contemporary methods of evaluation, utilizing assessments created in a manner divorced from curricular development. Tyler’s evaluation practices
followed directly from the development of the curriculum, as they were to assess the students’
ability to meet the defined curricular objective. The assessment followed the curriculum.
Current curriculum development is often structured to meet the demands of the mandated test,
and thus, the curriculum follows the assessment. Regarding concerns of instrument selection,
Tyler noted the importance of reliability and validity in drawing conclusions. Most importantly,
Tyler felt that assessments should be used for analysis, not summary description and labeling
(Tyler, 1949). He dismissed the notion that applying a single score to a student provided
adequate understanding of that student’s learning or any real assistance in improving curriculum.

Tyler’s work and its many interpretations and applications have impacted education and
schooling today. Of utmost importance is his determination that evaluation is central to
understanding schooling;

Finally, evaluation becomes one of the important ways of providing information about
the success of the school to the school’s clientele. Ultimately, schools need to be
appraised in terms of their effectiveness in attaining important objectives. This means
that ultimately evaluation results need to be translated in terms that will be
understandable to parents and the public generally. (Tyler, 1949, p. 125)

While his influence on the presence of testing has been noted, perhaps here he demonstrates not
only his influence on assessment but also on its application for uses in accountability. However,
missing from his analysis, and also from frequent analysis of testing, assessment, and
accountability today, is the question of when the evaluation of the objectives takes place.

Military Influence,

One of the important influences on educational objectives in American schools has been
the military and our sense of need for national defense. In the first half of the twentieth century,
while educators were involved in the development of assessment and evaluation of curriculum
and schooling, the nation found itself engaged in World Wars I and II. The use of testing was widespread throughout the military as personnel were categorized and sorted depending on their measured abilities in a variety of specializations. Military training requires a very specific skill set in mastery learning, as the same task must be accurately completed in a short amount of time even under great stress (Taubman, 2009). The connection to behaviorism is clear. As technological capabilities increased, the need for ever more efficient military training also increased. In the military, efficiency is critical, and task assignment and mastery learning are literally matters of life and death. In this model, all teaching and learning is standardized and mechanized, as a single method of accomplishing a task is clearly defined above any alternatives. Teachers must provide appropriate instruction, but there is essentially no creativity, flexibility, or individualization, and they are in many senses doing a job rather than working as professionals.

Technology and the standardization of information processing have been critical in the military and have impacted larger notions of education as the computational model of thinking has been embraced (Taubman, 2009). In fact, according to these theories of education, the emphasis on measurement of task performance has conceptualized knowledge as information and ability to demonstrate a behavior. This stands in contrast to earlier notions of the educated student as a participant in community and democracy, and excludes any alternative ways of knowing and displaying knowledge. The combination of the narrow conception of knowledge and task performance extend to abilities in problem solving. While this seemingly encompasses creativity of solution, the standardization of knowledge and routinization of task performance instead lead to a certain solution obtainable through the proper problem-solving techniques. Thus, knowledge here is totally reduced to that which is standardized and mechanical, with no variation or diversity.
Education in the Second Half of the Century

Science and Learning: The Impact of Sputnik and the Cold War.

With the Soviet Union beating the United States in the space race with the launch of Sputnik in 1957 and the fears of the Cold War on the heels of World Wars I and II, concerns of national defense were pressing on the minds of Americans well into the second half of the twentieth century (Taubman, 2009). Further, scientists and scholars were viewed with a newfound respect (Lagemann, 2000). This led to an interest in continuing the development of young scientists, which necessitated a rigorous education. The “Soviet success cast doubt on the quality of the American educational system” (Pinar et al., 1995, p. 154), and so the American public became concerned with developing and enhancing the teaching of science and technology in the schools.

Again, efforts were centered around defining behavioral objectives with measurable goals and outcomes (Pinar et al., 1995). The development of Bloom’s Taxonomy led to a more succinct breakdown of behavioral tasks into those achieveable on a hierarchy of domains, supporting the dissection of objectives into achievable, measurable tasks. This model of mastery learning supported the notion that anyone could learn anything, given proper instruction and time to learn (Pinar et al., 1995). This understanding of teaching and learning then further supports the importance of assessment, as if anyone can learn, surely any differences in achievement are based on the teaching and educational supports offered in the learning environment.

While much of this science and assessment movement was taking place within some domains of education and education research, there was continued debate and dissent, most notably amongst the curriculum theorists, who typically rejected the embrace of testing found in mainstream education. Instead, they focused on a crisis of meaning in education, finding that it
was too focused on achieving objectives that seemed divorced from the foundational goals of education that might not be measurable on the tests becoming so prevalent. Further, they were often found to be focusing on the larger societal issues impacting education, providing critiques based on gender, race, and political theory, amongst others (Pinar et al., 1995).

**Learning and Policy in the 1960s and Beyond.**

In alignment with the Civil Rights movement and the major, fundamental societal shifts of the second half of the twentieth century, education found itself undergoing extensive examination and review. A component of President Lyndon B. Johnson’s “War on Poverty,” the Elementary and Secondary Education Act of 1965 outlined specific areas of focus within education thought to bring about equality in an effort to close the achievement gap in newly desegregated schools (Meier & Wood, 2004; Pinar et al., 1995). The various titles of the legislation dealt with programs for educationally disadvantaged children, school libraries, educational innovation, and educational research, as well as funding for state educational departments (Pinar et al., 1995). Within the discourse of the Civil Rights Movement, educational equity and equality of opportunity to learn were important fundamental goals in need of attention. Further, as an element of Johnson’s war on poverty, the relationship between education and economic status is clearly recognized, a relationship that continues to exist and which informs much of the contemporary debate regarding school reform. Students attending high socioeconomic schools consistently outperform their peers at less affluent schools (Arbuthnot, 2011; Darling-Hammond, 2010; Ravitch, *The Death and Life of the Great American School System*, 2010). In working to address problems of poverty, statistically more prevalent
amongst people of color, and in connecting education to the solution, Johnson clearly saw the connection between the three, which is still relevant and problematic today.

**The Coleman Report.**

One of the mandates of the Civil Rights Act of 1964 was what came to be known as the Coleman Report, which was a study of equal educational opportunity (Lagemann, 2000). The purpose of the study was to survey equity in education. Typically, these surveys investigated inputs, such as teacher-student ratio, building quality, and library resources (Lagemann, 2000). However, the survey methods used by Coleman went beyond the traditional survey to include achievement testing data in addition to survey results and information regarding educational inputs. Coleman expected to find dramatic differences based on race. However, the results found essentially no interschool differences in achievement amongst students regardless of race, despite the fact that minority students’ scores were typically lower. Results, however, did vary based on student socioeconomic status and region, indicating that student background and socioeconomic status were barriers to achievement (Lagemann, 2000; Moss et al., 2008). Further, achievement gaps expanded as students aged, demonstrating that schools were ineffective in overcoming these barriers. While the findings are based on standardized achievement tests, which of course can be flawed, they still demonstrate the inequities present in education almost 50 years ago, wherein students whose families earn less money find themselves receiving an education not providing them with the adequate supports for measurable success. This achievement gap persists today. While the Coleman Report pointed to the important educational consequences of economic inequality, these factors have been largely ignored by policy makers. Instead, the focus is typically on teachers and in-school factors.
Given the impact of the Civil Rights Movement and the subsequent Women’s Rights Movement in America, the notion of schooling and education as political was fully understood and embraced in the 1970s (Pinar et al., 1995). Curriculum theorists found themselves playing the role of consultant in education (Pinar et al., 1995), a position that might seem versatile and flexible in a variety of contexts, but which also lacks a permanency and force necessary for informing and shaping curriculum directly. Thus, the field struggled to inform policy, a struggle that in many respects continues today. Educational policy had become distinctly reactive to social impulses rather than a force in shaping those impulses. The variety of important theoretical vantage points further complicates, while enriching, the discourse, again adding a level of complexity in the practical application of curriculum theory research.

A sense of crisis overwhelmed the discourse of the 1970s surrounding schooling as achievement gaps continued, test scores declined, and the notion that graduation standards had declined permeated the discussion (Lagemann, 2000). While panic struck those inflamed by the hegemony, many curricular theorists saw the political nature of schooling and its function as reproducing the economic structures of society. The aggressive participation of business and economic forces in shaping schooling had created the inequities demonstrated by the studies conducted in the 1960s. Notions of a “hidden curriculum,” the “unintended but quite real outcomes and features of the schooling process” (Pinar et al., 1995, p. 248), were applied to the reactive political impulses often dominating educational discourse, and education was often thought to be reproducing the very inequities it was hoping to alleviate. As the prevalence of testing and tracking rose, curriculum theorists seemed to anticipate the potential pitfalls of an

---

3 For a thorough discussion of the hidden curriculum, see Pinar et al., 1995.
education system too focused on testing, measurement, and classification, a climate reminiscent of the early days of eugenics. The hidden curriculum could constrain education into considering itself an easily definable and measureable institution.

**The Department of Education.**

The federal government first created an office dealing with education during the Civil War era, but the department moved around between various other departments, not finding a solid home nor substantive influence or interest, for many decades. With the rise in public interest in national security and the impact of science on that endeavor, however, the federal government did find itself engaged in educational policy. This was a major shift, as education had previously been thought to be the responsibility of the local community. This shift also led to a change in the role of teachers, whose individual importance and prestige continued to decline. In 1958, the National Defense Act greatly influenced funding for educational research. In particular, funding for the National Science Foundation was of great interest (Pinar et al., 1995). Recent technological advances as well as the Cold War had created a public interest in science education. Funding targeted scientific research, aligning with the focus on proof and measureable outcomes.

President Jimmy Carter extended the authority of the various government offices that had dealt with education by establishing the Department of Education and the cabinet-level authority it entailed. The move encountered much debate, as conservatives were concerned by federal involvement in education, which they saw as unconstitutional and in violation of states’ rights. Liberals, however, saw it as necessary in addressing the differences found between states as local interests overwhelmed national civil rights initiatives. The department continues to be
controversial and its role in education and policy is the focus of much debate. Its mandates are of course impactful, but their effectiveness is perhaps questionable.

**A Nation in Crisis**

The advances of the twentieth century brought about drastic changes in education and schooling, as the science of learning was utilized to determine set learning objectives for students that could be evaluated and measured through testing and assessment. While the inclusion of scientific reasoning in informing curricular decision making did encounter some objection, the model of results and accountability held strong, with the discourse now requiring the embracing of numerical data as proof of learning and teachers being judged by their students’ scores (Ravitch, 2010b; Taubman, 2009). Over the past generation, reforms have embarked on an even more accelerated path, as the test scores deemed problematic and only for use in very particular circumstances are used not only for student assessment, but for more dramatic high-stakes decisions, such as student promotion or retention, teacher salary or tenure, and school closure or takeover (Arbuthnot, 2011; Darling-Hammond, 2010; Ravitch, 2010b).

**A Nation at Risk.**

In 1983, *A Nation at Risk* was published, partly as an extension of the Coleman Report. This landmark study found that the state of education in America was dire (Lagemann, 2000). This and other educational studies of the early 1980s “accused American schools of decline, a lack of vision, even incompetence” (Pinar et al., 1995, p. 671). Much of the decline was based on threats of declining status in global economic competition, finding that schools should be more business-like, needing the business influence of measurable results. The need for regular
assessment of students was emphasized, a precursor to the mandated testing to come only a couple short decades later (Moss et al., 2008).

*A Nation at Risk* spawned various reform efforts, notably President George H.W. Bush’s *America 2000* initiative. With lofty goals, such as *all* children starting school ready to learn, a 90% high school graduation rate, first-place world rankings in math and science, and 100% literacy (Pinar et al., 1995), *America 2000* perhaps was never meant as a blueprint for schooling, but more of a vision of possibilities. Clearly a precursor to the No Child Left Behind legislation Bush’s son would enact from the same office, *America 2000* responded to its admirable goals through emphasizing a standardized curriculum that would be teacher-proof, a nod to the threat of incompetence indicated by *A Nation at Risk*. With the nation’s future in its hands, American education must be standardized to ensure success for all and not left in the hands of teachers, seen as untrustworthy and incompetent.

**No Child Left Behind.**

Under President George W. Bush, the reauthorization of the Elementary and Secondary Education Act took place in 2001, going into effect in 2002. As it has been widely known, the No Child Left Behind legislation was modeled after the Texas system that boasted so much success under Bush when he was governor of the state. While its actual success is debated and seems exaggerated, the model nonetheless came to dominate educational policy and discourse (Jesness, 2000). The law mandates annual standardized assessments for public school students in grades three through eight, as well as an additional assessment in grade 10 or 11. The results of the assessments must then be disaggregated by socioeconomic status, gender, race, ethnicity, disability, and English proficiency. However, no specific test of measurement is required.
The goals of the legislation are, as with much education reform, admirable. Bush sought to address the “soft bigotry of low expectations” (Bush, 2000), what he saw as the now assumed achievement gap leaving minority students behind their white counterparts. Through this legislation and the associated targets, incentives, and punishments, Americans could expect “higher-quality, more equitable, and more accountable public schools” (Meier & Wood, 2004, p. xi). Most importantly, “the broad goal of NCLB is to raise the achievement levels of all students, especially underperforming groups, and to close the achievement gap that parallels race and class distinctions” (Meier & Wood, 2004, p. 3). These levels of achievement could be reached through research-based educational practices, many of which emphasized the same assumptions of the learning sciences (Taubman, 2009). These higher achievement levels were not simply considered wishful thinking; rather, the law requires that all students reach proficiency on all mandated assessments by 2014, a deadline now fast approaching. While this is obviously a challenging benchmark, one might applaud efforts to push educators to raise the scores of all children. However, dissenters speak out about the inevitable failure of such standards, as the mandates also require standards be challenging. Thus, how is one to design an assessment that is both challenging and attainable by all students? (Ravitch, 2010b)

Perhaps more importantly, the focus on test scores shows a confusion of results with opportunity (Moss et al., 2008; Taubman, 2009). Student achievement continues to correlate most directly with socioeconomic status, and investigation of student test results divorced from the context surrounding their opportunity to engage in meaningful education is woefully devoid of meaning. Decades of research have demonstrated that students whose circumstances involve fewer educational resources and less funding for schooling perform more poorly than their more advantaged counterparts (Arbuthnot, 2011; Darling-Hammond, 2010; Lagemann, 2000; Pinar et
al., 1995). While teachers play an important role in providing instruction and opportunity to these students, it is optimistic to say the least to imagine that they can overcome all of these barriers and reach 100% proficiency.

Still, though, test scores, under No Child Left Behind, are utilized for high-stakes determinations of accountability, as schools are held responsible for making Annual Yearly Progress goals, raising student test scores each year. If they are able to make progress, they may be eligible for rewards. However, failing scores, and even more dangerous, repeated failing scores, will result in sanctions, threatening jobs and even the existence of the school. Takeovers by private charter school organizations are often one of the consequences, despite showing limited gains themselves (Ravitch, Who Kidnapped Superman?, 2011). These sanctions work to effectively dismantle the public school system as a whole, as teachers and schools are set up for failure in a system that requires unachievable results (Meier & Wood, 2004).

Many problems have been identified with the No Child Left Behind legislation, and one would be hard-pressed to find many full supporters at this point. It has been called “not a tool for solving a crisis in public education, but a tool for creating one” (Meier & Wood, 2004, p. 65), and critics from politicians and policy makers, to teachers, parents, and activists, and academics all find reason to challenge its policies (Taubman, 2009). In practice, the debate continues over what tests actually measure and whether or not they are accurate representations of student learning, while at the same time, states are found to be lowering their requirements for proficiency (Meier & Wood, 2004). Again, test scores are being confused with quality schooling, to the detriment of all involved in the actual educational experience of the classroom.
Testing, Accountability, and Contemporary Discourse

Recent reform efforts have not only embraced testing and, more specifically, high-stakes standardized assessments, they have actually framed the discourse to the extent that schooling is understood almost exclusively in terms of student performance. While assessment developed through a variety of means, testing characterized as the evaluation of student performance based on performance on a particular examination has come to define not only assessment but education as a whole. The results of the tests are now being used for purposes of accountability in a much more extensive manner than ever before, accelerating the audit culture that has become education in America. Even a supposedly more liberal administration has embraced these notions in the Race to the Top initiative. Further, the validity and appropriateness of testing are questions not often discussed enough.

Testing.

Beginning with Thorndike’s belief in the importance of numerical measurements for understanding and appreciating student learning, and the influence of the eugenics movement, standardized achievement tests were created and utilized on a large scale early in the twentieth century (Lagemann, 2000; Pinar et al., 1995; Winfield, 2007). An explosion in the development of tests was seen throughout the first few decades of the century, culminating in approximately 2600 tests by 1940 (Lagemann, 2000). Tyler was instrumental in the development of the National Assessment of Educational Progress (NAEP) test; while the test was seen in many ways as a step forward in assessing student achievement, criticism focused on the lack of practical recommendations for school improvement following its results (Lagemann, 2000). The test seemed to identify potential problems but had no suggestions as to how to address them. While
the interest in assessment was expanding, so too was the population of students, leading to increased interest in sorting students into groups theoretically most likely to find success as a cohesive unit within the classroom (Moss et al., 2008).

The military paralleled this interest in sorting and classification, and interest in testing for college readiness also increased, leading to the development of the Educational Testing Service in 1947, an influential party in testing (Lagemann, 2000; Taubman, 2009). Questions of achievement versus ability were teased out, the result being an enhanced focus on education as the means to increasing achievement regardless of ability (Lagemann, 2000). Quality instruction could lead to achievement, which could overcome any initial deficit due to ability. The measureable gains of this quality instruction then became the foundation of high-stakes consequences tied to student performance. Beginning in the 1970s, sanctions and rewards begin being tied to test scores, which then explodes with the implementation of No Child Left Behind (Darling-Hammond, 2010; Ravitch, 2010b; Taubman, 2009).

The ostensible goal of assessment is to inform educational practices (Moss et al., 2008). However, the variety of influences on its development and implementation impact and potentially distort these goals. We might typically assume that educational professionals and government officials would play a role in policy development, which is of course accurate. However, in addition to these obvious participants, major philanthropic organizations and corporate professionals also play a major role in educational policy creation (Taubman, 2009). While the motives of both groups may be pure and grounded in a fundamental belief in the importance of quality education for all citizens, we also would not be surprised to find that various entities bring with them the interests and biases of their financial ties. The discourse surrounding education has embraced the language of business, including the expectations of
gains in order to satisfy stakeholders. Further, the expectation that schools must prepare students to enter the global marketplace has dominated the movement towards aligning school curriculum with business standards. The influence is not always, nor frequently, explicit, and many question the notion of school as business (Taubman, 2009). Further complicating matters, many businesses and industries have found great economic success in the testing boom, even when the materials are not supported by the research supposedly required for implementation (Lagemann, 2000; Taubman, 2009).

The implementation of testing has clear challenges; further, the tests themselves are problematic in terms of their reliability and validity, as well as their decontextualized nature (Arbuthnot, 2011; Lagemann, 2000). For almost 50 years, they have been found to demonstrate an achievement gap that seems to measure family income moreso than student ability (Arbuthnot, 2011; Lagemann, 2000; Moss et al., 2008; Taubman, 2009). Attaching high-stakes consequences to the tests seems to actually suppress student performance, rather than motivating superior performance, as was hoped (Amrein-Beardsley et al., 2010; Paris, Lawton, Turner, & Roth, 1991; Taubman, 2009). Finally, the focus on student performance on a series of standardized assessments with high-stakes consequences drastically impacts curriculum, as teachers feel compelled to do whatever they can in order to drive their students’ scores upward. Moss et al. (2008) write:

Current assessments seem to focus primarily on coverage of subject matter content and basic skills. It may be that successful performance on these assessments also requires other aptitudes or capabilities, but if teachers and other educators who are held accountable for students’ success on these assessments think that their main focus is on discrete factors and skills, the tests may ‘drive’ instruction to concentrate on just those things. (p. ix)
The format and structure are clearly of critical importance, and so any assessments must be intentionally formulated to provoke the type of instruction and learning sought. Tests too focused on multiple choice questions about basic skills will shape curriculum to address those specific skills, as “testing shapes ideas of what counts as learning” (Moss et al., 2008, p. 23). Not surprisingly, too much focus on tests leads to classrooms in which “test preparation is the order of the day” (Meier & Wood, 2004, p. 39). Any assessments used in the classroom must be contextualized and applicable to the skills one hopes the students will learn. If the assessment measures only one type of learning, that demonstrated on a multiple choice test, only one type of learning is likely to occur. However, assessment of a variety of skills would by extension lead to a more diverse curriculum. Teachers often speak out against this narrowing of the curriculum, but still the focus on tests remains and teachers are attacked for “teaching to the test,” despite their frustrations that they have no other option.

**Accountability.**

The focus on accountability is a transformation within education that impacts everything (Taubman, 2009). Once teachers are measured by their students’ test scores and their impact on those numbers, their educational focus is forever realigned. Teachers resisted being measured by student performance early on, but their opinions were also highly valued and even considered central to curriculum development (Pinar et al., 1995). Contemporary discourse marginalizes teachers to the extent that, not only are their views often dismissed, they are labeled harshly as incompetent and lazy. While accountability proponents claim that it can hold teachers and schools responsible for their students’ performance, it seems that they too often fail to follow through with the appropriate funding and support needed to address student deficiencies. Thus,
typically accountability measures mistake measuring schools for fixing them (Meier & Wood, 2004). The ramifications for conflating teacher and student performance could range from fewer teachers being willing to work with high-needs and at-risk students to rampant cheating, which has already been occurring.

**Audit Culture.**

In Peter Taubman’s *Teaching by Numbers* (2009), he outlines the “audit culture” found in the discourse of education in America. Our need for data defines education and shapes our entire understanding of schooling and the educational enterprise (Taubman, 2009). Further, data provides fuel for a discourse that requires continuous improvement, as scores should be ever rising. The data then actually leads to the exclusion of the humans and human subjectivity, as all are defined by the data they create (Taubman, 2009). This discourse is predicated on an appreciation for economics and the marketplace, an appreciation that lends itself to an extention of those principles into essentially all areas of society, and education in particular. This neoliberal impulse utilizes the corporate model, traditionally measured by financial gain, to measure achievement by assigning a numerical value of worth. The free market is applied under the auspices that what is good for corporations is good for education and vice versa. Corporations need the highly-trained employees that schools provide, and schools need corporations to hire their graduates, thus creating a friendly partnership. However, integrating corporate leaders into the creation of educational curriculum and assessment assumes that these corporate leaders have an understanding of the educational practices that will most benefit their future employees. Instead, we might wonder if perhaps the professional educators are more qualified to make those determinations. Education and its assessment has proven to be a
lucrative business, however, and so corporate influence is not likely to diminish anytime soon (Taubman, 2009).

The control exerted by corporate leaders in conjunction with policy officials often exceeds the influence of teachers in making critical educational decisions (Taubman, 2009). The control of those in power is seen both in their definition of educational policy as well as through their surveillance from afar. The need for this surveillance to be easily quantified restricts the possible use of more comprehensive authentic assessments. Further, as economic efficiency is held in the highest regard, less money is allocated for education, which leads to even fewer resources allocated for meaningful assessment; “as resources devoted to education decline, it is even less likely authentic assessments or differentiated assessments will be used, given the time and personnel required for their implementation” (Taubman, 2009, p. 22).

Learning is thus defined exclusively as performance on a single high-stakes standardized assessment and the promises of the learning sciences are once more embraced. Paired with the standardized mechanization of teaching, the learning sciences and these assessments “promised and promise certainty, status, and a defense against the turbulent unpredictability of the classroom. They claim that if they are followed, all students will succeed, and social problems will be solved” (Taubman, 2009, p. 182). If students can excel on the tests, we will know that they have learned and are therefore educated and capable of productively engaging in society. The role of teachers is emphasized as responsible for this performance, but also diminished and deprofessionalized, as they need only follow a script in order to ensure learning. Learning, again, equates to performance on a test, an abstraction of education that in contemporary discourse actually harkens back to the educational debates going on a century ago.
The Role of Teachers.

Teachers are frequently noted to be the most important factor in student learning. A more accurate description is that they are the most important in-school factor in student learning, as students’ backgrounds and family and socioeconomic environments are strong influences on achievement (Darling-Hammond, “Securing the Right to Learn,” 2006). Still, though, even when described with the glowing adoration often found in media portrayals of successful teachers, their success is determined by their students’ ability to succeed on a test.

Still, the reforms and the discourse of crisis and failure have shocked educators into compliance with the demands of education today (Taubman, 2009). Their role has then become diminished to only teaching students to perform on a single test. Further, they are restricted to standardization of curriculum geared towards test performance, ignoring the nuances of the classroom (Moss et al., 2008; Taubman, 2009). The forgotten component is the student, whose humanity and singularity of experience has been all but dismissed. Further complicating matters, there is no room for teachers to assess and address gaps in individual students’ background knowledge, which might otherwise inform instruction in meaningful ways (Taubman, 2009).

With the challenges and frustrations of education thus fully apparent, Taubman (2009) postulates theories as to why teachers accommodate the requirements of the reform initiatives, while frequently simultaneously disagreeing with their premises and implementation. The four components of his argument stem from psychoanalysis and are one interpretation of how the discourse of testing and reform policies impact teachers. Taubman describes a sense of fear, shame, fantasy, and guilt that connects teachers to their profession despite its frustrations and challenges. He claims that teachers fear a lack of control and also fear failure, especially as their performance is linked to their students’ test scores. Poor test scores can then lead to a lack of
resources, exacerbating teachers’ fears. The discourse of failure leads to a sense of shame for teachers, as they are frequently criticized as incompetent and lazy. However, they are also often hailed as heroes in the lives of children, leading to a fantasy about the possibilities of the future. Finally, the guilt teachers feel due to the failed mission of schooling may lead them to go along with reform efforts. Public schools were supposed to be the ultimate mechanism of social justice and were to fix the social ills of segregation and poverty. However, schools are now more segregated than at any time since the passage of Brown vs. the Board of Education (Ravitch, Who Kidnapped Superman?, 2011; Taubman, 2009), and public schools seem to have failed in their mission of equity. Teachers thus find themselves at a crossroads, frustrated by reform policies but also inclined to keep working towards the educational goals that brought them into teaching in the first place.

The evolution of educational thought since the end of the Civil War in the United States has at times embraced the humanity of students but more commonly their measurement. In order to know that learning has occurred, the loudest voices in the discourse profess that students must demonstrate some proof of that learning via testing. Education is both elevated and reduced to the status of science, as it is measurable through scientific means while teachers are dismissed as the mere relay mechanism necessary for the instruction of the standarized curriculum. However, we must remember to embrace the nuance of education (Taubman, 2009). If education consists of that which can be measured, then we must ask the question Taubman asks: “What if the obsession with learning keeps us on track but also keeps us from being educated?” (2009, p. 195). We must remember that schools operate within context, and assessments should do so as well.
Testing and the importance of student success on high-stakes standardized assessments impact teachers and their instructional practices. The purpose of this study is to survey teachers about their experiences with high-stakes standardized tests and the influence of these assessments on their schools and teaching. The primary research question asks how teachers feel about testing, its impact on their instructional time, the pressure it puts on teachers, and the support the school provides.

**Contemporary Related Studies**

As assessment is currently most commonly being defined as standardized testing, it is important to understand how testing operates in the classroom. Good, valid tests are needed if their results are going to be representative of student learning and helpful for teachers as they develop instruction to meet student need. While many teachers are in agreement that evaluation and assessment are beneficial, and even crucial, to effective instruction and schooling, high-stakes standardized assessments as currently implemented create challenges for teachers and students. The emphasis on accountability based on student performance on these tests creates a significant amount of pressure for teachers, who often find themselves “teaching to the test” and even resorting to cheating to ensure adequate scores. The changes in instruction may result in teachers feeling they are not addressing student need. Also related to the pressure and curricular changes is the role of principals and administrators, who may help alleviate some of the teachers’ concerns, or, conversely, exacerbate the problems.
Testing, Measurement, and Application in Schools

As standardized tests are currently the primary method of assessment used in schools, in order to accurately assess students, appropriate tests must be used and their results must be interpreted correctly. Tests should be designed to ensure validity, but this can be problematic as tests are not always used as intended by their creators. This can lead to unfortunate results as some groups of students struggle with the tests. Assessments can be used effectively, but there are challenges associated with doing so and also with their role in accountability systems.

Arbuthnot (2011) provides an overview of the process involved in developing tests, including validity and test fairness, and test performance patterns in *Filling in the Blanks: Understanding Standardized Testing and the Black-White Achievement Gap*. As testing use increases, validity is of the utmost importance. Validity is the extent to which the “evidence supports the intended interpretation of test scores for the proposed purpose” (Arbuthnot, 2011, p. 10). Thus, does the test actually measure that which it is intended to measure? Arbuthnot gives an overview of the Black-White achievement gap, noting that White students typically outperform their Black colleagues, leading to questions regarding the validity of the tests. Perhaps the tests are not designed and utilized appropriately, leading to differing performances by students of different races. Thus, testing is a complicated matter with important implications for students, as they are labeled by their scores, which might not actually represent their learning or ability.

In *Spotlight on High-Stakes Testing* (2003), the contributors provide a policy analysis by reviewing the complications and challenges of test-based reforms with high-stakes consequences. Concerned with the focus on testing, the authors address the potential benefits of high-stakes tests and how educators might find opportunity in testing, while also weighing the
potential for a negative impact on schooling and curriculum. The authors find that educators
should advocate for appropriate and valid assessments that provide useful data and feedback
while avoiding the potential pitfalls of teachers feeling the need to teach to the test or adjust
instructional strategies toward drilling test concepts. Appropriate assessments will have a
reasonable number of well-defined objectives that provide insights to teachers regarding their
students’ mastery. With appropriate assessments and collaboration regarding effective
instructional strategies, testing can actually be beneficial for teachers and students. They present
a common theme, which is that test-based accountability efforts can be positive for schools if
handled appropriately through solid assessments and strong instruction.

In “Controversies of Standardized Assessment in School Accountability Reform: A
also provide a policy analysis and further elaborate on the arguments both for and against testing
in schools. They divide the issue into four categories: assessment-driven reform, standards-
based assessment, assessment-centered accountability, and high-stakes consequences.
Essentially, assessment-driven reform refers to the current focus on attempting to improve
schools through the implementation of standardized assessments. These standards-based
assessments are tests designed to measure students’ achievement regarding prescribed standards.
The assessment results are then used to hold teachers and schools accountable for student
learning and performance, and high-stakes consequences in the form of sanctions and rewards
are attached to those test results. The researchers find that there does seem to be evidence that
American students are lagging behind their international peers and assessment can be valuable,
although the tests cannot guarantee improvement alone. Setting standards for the assessments
can be tricky, as many people argue that there should be certain high standards for students’
learning and education, while others find standards to limit intellectual freedom, diversity, autonomy, and teacher empowerment. Situating accountability on the results of these assessments is also a topic of much debate, as not only are the tests questioned, but also the responsibility of the teacher for student improvement given the contextual challenges students face, such as socioeconomic status and family support and structure. However, the tests can be used to ensure high standards and coherent curricula. Finally, some argue that high-stakes consequences are necessary for reform efforts to be taken seriously. Others, though, find that the consequences are problematic in that they do not appreciate the complexity of learning and can lead to unfortunate results, such as students leaving school prior to graduation and an overemphasis on test preparation. Through examining the pro and con arguments for each of these issues, Wang, Beckett, and Brown (2006) provide a comprehensive overview of the various challenges associated with the current standards-based reform efforts.

The debate regarding testing is thus rich with strong potential for positive impact on education but also significant challenges in implementation. Most seem to agree that tests must be valid in order to provide an accurate measure of student learning and then must be implemented effectively in order to avoid the pitfalls often associated with these assessments, such as high student dropout rates, narrowed curriculum, and the loss of teacher autonomy and professionalism.

**Teachers’ Opinions of Testing**

As the primary stakeholders impacted most directly on a daily basis, teachers have developed diverse opinions about testing and its impact on their teaching and classrooms. As the results of the studies described below demonstrate, while teachers tend to be supportive of
evaluation and believe it is important for effective schooling, they are typically less enamored by the testing as implemented due to targeted policies placing a strong emphasis on high-stakes standardized assessments.

By surveying 558 teachers, James and Pedder (2006) examined teachers’ attitudes towards assessment practices and how the values espoused by testing and assessment might align or challenge their own values regarding education. The participants of the study were teachers working in 32 schools in England. While policy requirements in England differ from those in the United States, the emphasis on standardized assessments used for purposes of accountability is similar and the results are relevant to studying testing policy in America as well. The researchers examined three dimensions of assessment: making learning explicit, promoting learning autonomy, and performance orientation, which they define as the fundamental principles for using assessment in the classroom. Making learning explicit entails “eliciting, clarifying and responding to evidence of learning; working with students to develop a positive learning orientation” (James & Pedder, 2006, p. 122). This facet involves demystifying the learning process and goals for students in order to create a more positive attitude towards school and learning. A focus on promoting learning autonomy allows students independence over their learning objectives and assessment. Performance orientation refers to assisting students in complying with the performance goals as defined by the curriculum. Through blending the three, teachers can align assessment with curricular goals. While the first two are formative and more easily aligned with assessment for the sake of improving learning, the third is more directly assessment of learning, and therefore more summative. The researchers found that teachers described their own values as being in alignment with making learning explicit and promoting learning autonomy, and while they understood the importance of performance orientation in the
culture of assessment, it less directly aligned with their own values of teaching. Its emphasis on test results was found to create tension for teachers. The authors found that teachers are committed to assessment for the sake of learning and that this type of assessment aligns with their personal values and impetus for teaching, and that they face a challenge in aligning their values with assessment of learning, however important it is for policy makers.

Monsaas and Engelhard (1990) investigated how teachers’ attitudes towards testing impact their classroom practices. The authors used the Testing Practices Instrument survey and sampled 186 teachers in Georgia. The participants were all taking graduate courses at a college and are thus a somewhat selective group of teachers working to further their skills or perhaps qualify for additional compensation by earning class credit or an advanced degree. The researchers examined behavior, attitude, and pressure to investigate how teachers’ attitudes towards testing and the pressure they feel impacts their testing practices behavior, such as test preparation practices. They found that teachers engaged in more testing practices aimed at raising test scores when they felt pressure, although it was not a strong correlation. Instead, their attitudes towards testing were better predictors of testing practices. They also found that when teachers believed particular practices were essentially cheating the system, and geared only towards raising test scores and not at an increase in learning and skill development, they were less likely to engage in those practices. Teachers with higher enrollment of students from lower socioeconomic groups were more likely to engage in test preparation activities, and teachers of earlier grades were also more likely to engage in such practices. Thus, the researchers found that testing did impact teaching practices, although mediated by teachers’ attitudes towards particular test preparation practices and student characteristics.
Craig (2010) finds that teachers believe that evaluation-based school reforms in America have actually overwhelmed their own intents and desires as practitioners. In a qualitative study using narrative inquiry, Craig developed relationships with teachers and principals in five schools in the United States encompassing all grade levels, including one high school, two middle schools, and two elementary schools. These relationships allowed the researcher to learn the stories of their experiences with testing and assessment, which are documented as narratives. Craig (2010) examined teachers’ perspectives on evaluation and the meanings they make of their work, looking particularly at their experiences with outside evaluators. Craig believes these outside evaluators do not take teacher experience and meaning-making sufficiently into account and do not appreciate who the teachers are who will be implementing testing and reform policies. While the divide between theory and practice has been frequently investigated, Craig expands that and defines the “theory—practice—reform split” (p. 1298), noting that one-size-fits-all solutions do not exist and reforms should be situated within the specific school context.

Not all teachers feel testing and high-stakes assessment practices are negatively impacting their teaching. Buck, Ritter, Jensen, and Rose (2010) further investigated teachers’ experiences of testing, conducting focus groups, finding several positive themes associated with test-based school reform. The researchers interviewed 42 teachers and principals in five schools in Arkansas in focus groups of eight to ten participants each. The teachers believed testing provided useful data that helped create a plan for instruction and that tests do not necessarily hinder creativity or collaboration, and actually believed that accountability is useful. The teachers who participated in this survey were not working in particularly low-achieving schools, and thus the pressures they felt might be less than some others, but they still represent a group of teachers who find testing and accountability helpful and beneficial for instruction and schooling.
While testing can provide guidance regarding student struggles that can inform curricular decisions, teachers’ attitudes towards the tests and accountability measures can impact their teaching. Formative assessments are particularly useful, but teachers are wary of guiding instruction too directly towards testing in a manner that could verge on cheating. When assessments are in alignment with teacher beliefs and values regarding education, teachers actually see accountability policies as helpful in improving their instruction.

**Pressure on Teachers**

The application of high-stakes consequences based on student performance on standardized assessments places strong pressure on teachers to ensure adequate, if not exceptional, performance. This pressure leads to changes in curriculum in many cases, especially in schools whose students typically do not perform well, and sometimes even leads teachers to cheat in order to raise their students’ scores.

In “State Standardized Testing Programs: Their Effects on Teachers and Students” (2007), Moon, Brighton, Jarvis, and Hall report their findings from a survey and qualitative study of teachers and students regarding their views on the impact of state testing programs. Nationally, 1289 elementary school teachers as well as 415 middle school and 393 high school teachers completed the survey, and small (3-5 participants) student and teacher focus groups were also formed and analyzed. The researchers found that teachers feel a tremendous amount of pressure and that this pressure leads to a shift in curriculum and instruction to drill and practice activities. Further, the pressure is felt even more strongly in disadvantaged schools and the drill and practice instruction is even more prevalent.
Perreault (2000) also conducted teacher focus groups to ascertain teachers’ opinions of the impact of mandated testing as well as the legitimacy of these programs and reported the findings in “The Classroom Impact of High-Stress Testing.” Seven to nine teachers in the same state from schools both highly successful and those struggling on mandated assessments participated in each of the focus groups. The teachers voiced concerns about the tests, noting that they felt constant pressure to make sure their students fared well on the assessments. Further, they felt that they were pressured to develop their curriculum around test requirements and “teach to the test.” The pressures made them feel “‘defeated,’ ‘powerless,’ and ‘unsure if they were doing the right things’ to help students succeed” (Perreault, 2000, p. 4). These teachers clearly voiced concerns regarding the pressures associated with these assessments.

In “Test Anxiety: A Multifaceted Study on the Perceptions of Teachers, Principals, Counselors, Students and Parents” (2005), Mulvenon, Stegman, and Ritter surveyed a variety of stakeholders and linked their results to student performance on national and state assessments. Surveys designed for each particular group were administered to teachers, principals, counselors, students, and parents. The survey findings of the teacher portion are most relevant for this study. 141 teachers participated in the study. While most stakeholders did not support claims of the “dangers” of test-based school reform, teachers did demonstrate concerns and anxiety due to the assessments. The researchers found that student achievement and teacher attitude are not related and students are not negatively impacted by teachers who have concerns about testing. However, teachers did report worry that test information would be used against them unfairly on performance evaluations and acknowledged that they occasionally break testing protocol to assist students. The anxiety teachers feel leads to the potential for cheating, but fortunately does not otherwise impact student performance.
Amrein-Beardsley, Berliner, and Rideau (2010) further explore teachers’ propensity to cheat in “Cheating in the First, Second, and Third Degree: Educators’ Responses to High-Stakes Testing.” Through a survey of 3085 teachers in Arizona, the researchers found that over 50% of the respondents said they had known of colleagues who had cheated, and more than 50% of them further reported having done so themselves. This large of a group is of significant concern. The authors found that, “when pressured to do well on high-stakes tests, educators engage in quite clever practices, largely to protect themselves and their students” (Amrein-Beardsley et al., 2010, p. 25). The authors remark that teachers are typically believed to be of high moral character and that most people are shocked by the notion that they might cheat, but the high-stakes testing policies produce such extreme pressure on teachers that they at times find themselves engaging in these inappropriate activities.

As these studies found, many teachers feel intense pressure because of high-stakes assessment policies. This pressure causes anxiety for the teachers and has implications for schooling. Many schools make curriculum revisions in order to better align their instruction with what is tested on the assessments and teachers find themselves “teaching to the test.” Further, a potentially high number of teachers even resort to cheating in order to raise test scores. These actions clearly indicate the extreme pressure felt by teachers and the consequences of that pressure on the classroom.

**Instructional Time**

Many researchers have examined the notion of teaching to the test more explicitly, as it has become a widely discussed trend that frustrates teachers and is viewed as problematic even to those outside the institution. The curricular shifts are particularly important in middle school,
a developmental time for students that finds them facing many social, physical, and emotional changes, and middle school curriculum has traditionally embraced these shifts with collaborative and student-centered work. However, the testing mandates have led to changes in this focus, with instruction often more teacher-driven and focused specifically on testing.

Musoleno and White (2010) conducted a survey of 148 educators in Pennsylvania to assess the impact of high-stakes standardized assessments on instructional time, finding that instructional practices have been altered to include additional time for test preparation, and more time is spent on tested subject areas to the detriment of non-tested subjects. They looked specifically at middle school students, who have particular developmental needs due to their adolescence. Stating that their physical development and social needs are best served by instructional practices allowing for interaction and cooperative learning, the researchers found that teachers were frustrated by the impact of NCLB and the time focused on test preparation and testing. They felt that they “lost opportunities to be creative and flexible and are currently employing more directed, teacher-led instruction” (Musoleno & White, 2010, p. 9), rather than utilizing practices more aligned to their students’ particular developmental needs. Further, the authors noted that tested subject areas were allotted increased instructional time to the detriment of non-tested areas.

Faulkner and Cook (2006) also investigate the impact of testing on instructional practices, finding that the assessments have led to more teacher-focused instructional methods. They used the Middle School Concept Implementation Survey and collected responses from 216 middle school personnel in Northern Kentucky. They find that “though teachers acknowledge the importance of including active and student-centered strategies on a consistent basis, the state tests seem to drive the curriculum and warrant more teacher-focused instructional methods—
lecture, worksheets, and whole-class discussion” (Faulkner & Cook, 2006, p. 1). The researchers surveyed teachers and found that they feel the tests lead them to teach to the test and focus on covering the tested material instead of in-depth study.

In “What are NCLB’s Instructional Costs?” (2006), Zellmer, Frontier, and Pheifer explore the impact of test-based reform efforts on not only classroom instructional time, but on additional resources spent on logistical planning and testing administration. The researchers sent a survey to an administrator in every school district in Wisconsin with a 40% response rate of 171 respondents. Through surveying these education leaders who work with the tests, the authors found that there were opportunity costs of logistical preparation, test administration, and loss of instructional time and services, as teachers’ time is spent on test administration rather than instruction. They also noted financial resources dedicated to testing rather than instruction, and found that respondents also noted a narrowing of the curriculum.

In “Constrained Professionalism: Dilemmas of Teaching in the Face of Test-Based Accountability” (2009), Wills and Sandholtz conducted a qualitative study of an experienced teacher’s encounters with the accountability demands of NCLB and how the focus on assessments impacted her instructional time, finding that despite a supportive administration confident in her abilities, the teacher found herself devoting more time to tested areas and providing instruction in other areas that she did not feel best met her students’ needs. The study consisted of a case study of a single fifth-grade teacher in a rural school in Southern California serving many first-generation United States citizens, English language learners, and recipients of free- and reduced-price school lunches. The school had been low-performing but was also experiencing improvement. The researchers found that the testing led to standardization of instructional practices that did not respect teachers as professionals.
As these studies demonstrate, the impact of NCLB and mandated high-stakes assessments have drastically influenced instructional time and teacher autonomy. Many teachers find themselves following prescribed methods of instruction, essentially teaching to the test. Time spent on tested subjects detracts from non-tested areas. Despite the unique needs of middle school students, many schools’ curriculum is reorganized to address the needs of the testing situation.

**School Support**

Any school improvement policies impact not only teachers and students in the classroom, but also the school principals and administrators who must work to implement policies and improve performance school wide. School leaders and their leadership styles have been studied to examine what types of leadership and support might prove most effective in improving schools. Principals and administrators set the direction for their schools, including setting expectations for their staff and providing the supports they deem necessary for helping their teachers achieve those expectations.

Sunderman, Orfield, and Kim (2006) examine the important implications for the role of principals due to the No Child Left Behind legislation and accountability-based reform efforts in “The Principals Denied by NCLB Are Central to Visionary School Reform.” They surveyed teachers in two urban school districts in Fresno, California, and Richmond, Virginia, serving many low-income and minority students. The schools were identified as needing improvement and also as making adequate progress towards that improvement. The authors outline the role of school administrators in addressing the mandates as well as the challenges the position faces. Noting that lasting effective reforms require support and resources, the researchers surveyed the
teachers to assess their views of NCLB. They found that teachers believed that standards are important, but that rewards for strong performance would be more beneficial than the sanctions applied under NCLB. Further, they found that teachers believed that strong leaders are essential for reform and that reform is collective as opposed to individual. Principals should carefully consider the impact of reform efforts on instruction in their schools and should provide encouragement for practices that are working. Also, principals should emphasize the use of assessments for diagnostic purposes and ensure that testing does not take excessive time away from other educational endeavors. Finally, principals are crucial to creating a supportive environment that will encourage teachers to remain at schools, decreasing turnover, which can be problematic for sustained reform efforts.

Reitzug, West, and Angel (2008) further examine principals’ own understanding of their role within current education reform efforts in “Conceptualizing Instructional Leadership.” The authors classified their opinions on their roles as leaders into four categories and examined the possibilities and limitations of each. Through in-depth interviews with 20 principals, the researchers determined there were four conceptions of instructional leadership: relational, linear, organic, and prophetic. Relational leadership focuses on developing relationships with students and faculty to increase motivation, efficacy, and pride. Linear leadership practices assume “systems can be designed so that one action, process, structure, or intervention will lead to a subsequent desired outcome, which will then lead to the next desired outcome and so on down a causal chain” (Reitzug et al., 2008, p. 699). Through careful monitoring, these leaders believe that a precise system can be followed to lead to the educational goal. This form of leadership is directly responsive to the standards movement. Organic leadership practices are based on the notion of the school as a sort of living organism, with instructional components linked to the
larger school and societal environment. Prophetic leadership is based on the idea that school leaders are answering a call to be leaders and focuses on strong educational purposes rather than simply test scores. The authors believe school leadership is about skill but also about purpose and reject the linear model, while supporting both the relational and organic methods as most effective in creating a strong schooling environment.

In “Leadership Styles and High-Stakes Testing: Principals Make a Difference” (2002), DeMoss investigates how principals leadership philosophies impacted their approaches to their position in relation to the mandates of standardized assessments. By examining case studies at eight paired elementary schools in Chicago comprised of observations and interviews with teachers and principals, the author found that “the ways principals framed how their schools would respond to the testing environment was responsible for schools’ test performance” (DeMoss, 2002, p. 113). The principals studied varied in their approaches, from creating an exclusive environment with only students and teachers who would fit a rigid, prescribed approach, to treating teachers as professionals who are part of a structure working towards continuous improvement and professional growth, to a focus on personal relationships. DeMoss found that schools without a strong instructional focus were less successful than those with a well-defined approach developed through teacher and administrator collaboration.

In “Does School Leadership Matter in the NCLB Era?” (2011), Singh and Al-Fadhli examine the impact of administrators’ efforts to meet the requirements of NCLB, including not only student achievement, but also issues of funding, teacher support, and parental involvement. They analyzed standardized test scores and interviews with school leaders, both superintendents and principals, in Mississippi. The authors believe administrators must be “knowledgeable about assessment and skilled in using data to make instructional decisions for their respective schools
and districts” (Singh & Al-Fadhli, 2011, p. 752). They found that administrators were dissatisfied with the financial support provided, and also that test-taking strategies and parental involvement were felt to be important. Overall, they believe that a synergy through collaboration, development, communication, support, and shared goals is critical for change.

Each of these studies examined the role of administrators in school reform. The common theme that emerged most dominantly is that teachers and administrators must work together in a collaborative and professional manner in order to develop appropriate school goals and strategies for reaching those goals. Principals and administrators are key in creating an environment that helps teachers feel respected and autonomous while striving towards the common goal through the methods developed by the collective. Further, all must be knowledgeable about testing, data, and reform efforts, and how each informs the others. The role of principals and administrators thus cannot be taken lightly, as they are the leaders who shape the schooling environment, for better or for worse.

**Conclusion**

As these findings indicate, the implementation of school reform policies based on high-stakes standardized assessments is dramatically impacting teachers and schools. Concerns over test validity are somewhat overshadowed by the concerns teachers feel in the use of the test results, which they find to create tension and pressure in the classroom, leading to altered instructional practices and sometimes even unethical behavior. While supportive and collaborative leadership can help mitigate some of these factors, the fact remains that the emphasis on testing is having significant consequences on teaching and learning. As the policies of No Child Left Behind have now been implemented for a full decade, and the new Race to the
Top competition has extended the focus on test scores and school accountability, understanding teachers’ perspectives on how these policies are influencing their teaching is critical to appreciating their true impact. Contemporary discourse now not only focuses on “failing” schools based on these scores, including disparaging comments regarding the teachers working with students, but is actually now moving towards discussions regarding the abolishment of public education as a whole. Thus, new research investigating teachers’ experiences of these policies is needed in order to inform the debate and provide insights from those most directly involved in educating our children, their teachers.
Chapter 3
Methods

Procedures

It was determined that survey research methods would be utilized for this study to gather information from the target population of teachers. Surveys are useful for gathering participant opinion and understanding characteristics of a population (Fowler, 2009; Fraenkel & Wallen, 2009). As teachers have direct experience with high-stakes standardized assessments, gathering their opinions is valuable in understanding the impact of these mandates on teachers and their classrooms.

In order to investigate teachers’ experiences of the impact of high-stakes standardized assessments on their teaching, an adaptation of the National Center for Research on Evaluation, Standards, and Student Testing (CRESST) survey of the effects of standardized testing on teachers and learning was used. The variables examined in the survey are: “accountability pressure, teacher attention to testing in instructional planning and delivery, time spent on test preparation, teachers’ sense of professional pride, and general attitudes teachers hold about the fairness and utility of testing” (Herman & Golan, 1991). While many surveys were considered, the CRESST survey was the most comprehensive survey including the combination of constructs that best addressed the areas of concern highlighted by the review of literature.

The survey was sent to teachers at a public middle school in Baton Rouge, Louisiana. Louisiana has utilized the Louisiana Educational Assessment Program (LEAP) test in math and language arts since 1999, adding science and social studies in 2000 (The 3 P's of testing: Plan, prepare, pass!, 2009). In East Baton Rouge Parish, the schools have performed below state levels, and at the 8th grade level, most relevant to this population of teachers, the district scores have been improving in all subjects except math, which has been fairly flat since 2006 (EBRPSS
Committee of the Whole, 2011). Recently, major legislation has been debated and passed in the state legislature that will drastically impact public school teachers, from new teacher tenure and compensation regulations to an increase in charter schools and public funding used to pay private school tuition (Barrow, 2012). These conditions make Baton Rouge public school teachers a population worth surveying, as their experiences directly align with much of the discourse surrounding testing across the country. Teachers were surveyed in an attempt to learn about their experiences of working with high-stakes standardized assessments and their attitudes towards testing, its impact on their instructional time, the pressure they feel, and the support they experience from their administration. The survey was sent to 105 teachers at the school and 15 responded, a response rate of 14.28%.

Participants

The online survey was emailed to 105 teachers at a public middle school, grades 6-8, in Baton Rouge, Louisiana. The school tends to outperform the district but has lagged behind the state the past two years, after outperforming the state averages for several years prior. The school earned a 2010-2011 school performance score (SPS) of 90.7. The state of Louisiana has set a target SPS of 120.0 for all schools by 2014. They received a “C” letter grade, with 62% of students performing at or above grade level. The school had a total enrollment of 927, with 84.4% of students eligible for federal free- and reduced-price meal programs, a commonly-used indicator of student socioeconomic status. The school has a student population that is 89.9% minority and 8.8% of the students have disabilities.

The school also serves as a gifted magnet school and some of the teachers work with students classified as gifted and talented. Students participating in the gifted program take
classes targeted directly towards their needs and only participate with the general student body in elective classes. The talented program includes special art and drama classes for its students. Students in the gifted and talented programs might differ in their experiences of high-stakes testing from the general population, potentially affecting their teachers’ perspectives as well.

The teachers are 76% female (n=80) and 24% male (n=25), and 60% white (n=64) and 40% non-white (n=41). A breakdown of teacher race and ethnicity was not available other than a classification between white and non-white teachers. The respondents somewhat resembled the teacher population with regards to gender, but there were fewer non-white respondents than white respondents when compared to the racial breakdown of the school. Subgroups were examined in regards to gender, educational training, teaching experience, subject area focus, and work with gifted and talented students.

Table 1: Participant Demographics

<table>
<thead>
<tr>
<th>Demographic</th>
<th>Population</th>
<th></th>
<th>Sample</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
<td>%</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>25</td>
<td>23.8%</td>
<td>5</td>
<td>33.3%</td>
</tr>
<tr>
<td>Female</td>
<td>80</td>
<td>76.2%</td>
<td>10</td>
<td>66.7%</td>
</tr>
<tr>
<td>Race/Ethnicity</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White or Caucasian</td>
<td>64</td>
<td>61.0%</td>
<td>14</td>
<td>93.3%</td>
</tr>
<tr>
<td>Other</td>
<td>41</td>
<td>39.0%</td>
<td>1</td>
<td>6.7%</td>
</tr>
</tbody>
</table>
Table 2: Sample Subgroup Demographics

<table>
<thead>
<tr>
<th>Demographic</th>
<th>Sample</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$n$</td>
<td>%</td>
</tr>
<tr>
<td>Training</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bachelor’s Degree</td>
<td>8</td>
<td>53.3%</td>
</tr>
<tr>
<td>Advanced Degree</td>
<td>7</td>
<td>46.7%</td>
</tr>
<tr>
<td>Experience</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Novice (10 years or less teaching)</td>
<td>9</td>
<td>60.0%</td>
</tr>
<tr>
<td>Veteran (more than 10 years teaching)</td>
<td>6</td>
<td>40.0%</td>
</tr>
<tr>
<td>Subject Area Focus</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tested Subjects</td>
<td>9</td>
<td>60.0%</td>
</tr>
<tr>
<td>Non-Tested Subjects</td>
<td>6</td>
<td>40.0%</td>
</tr>
<tr>
<td>Gifted/Talented Student Percentage</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0%-50% of Students Classified as Gifted</td>
<td>10</td>
<td>66.7%</td>
</tr>
<tr>
<td>51%-100% of Students Classified as Gifted</td>
<td>5</td>
<td>33.3%</td>
</tr>
</tbody>
</table>

Instruments

An adaptation of the CRESST survey of the effects of standardized testing on teachers and learning was utilized to investigate teachers’ opinions of the impact of high-stakes standardized assessments on their teaching. The survey was developed to assess the pressure teachers feel to meet mandated goals on high-stakes standardized assessments, the impact of these tests on their instructional time, the school’s focus on test results, and the teachers’ feelings regarding the helpfulness and fairness of the tests. The authors of the survey found that teachers do feel substantial pressure to improve their students’ test scores and that the testing does impact their classroom planning and instructional time. However, the teachers they surveyed did not believe that testing was helpful in clarifying learning goals, giving helpful feedback, or helping schools improve (Herman & Golan, 1991). In order to focus the research questions and limit
participant requirements, for this study, the following areas were included: pressure, instructional
time, school support, and attitudes regarding testing.

The survey used in this study was somewhat abridged for the sake of brevity, and the
selected survey questions focused on teachers’ feelings of pressure from a variety of groups (i.e.
“I feel pressure from my principal to improve my students’ standardized test scores,” and “I feel
pressure from parents to improve my students’ standardized test scores”), its impact on their
instructional time (i.e. “I adjust the sequence of my curriculum based on what’s included in the
standardized test,” and “I spend time giving students practice in the kinds of item formats that
are on the standardized test”), the support they feel from their school (i.e. “Programmatic efforts
to improve student learning are present in my school,” and “My school administration discusses
ways to improve standardized test scores”), and their attitudes towards testing (i.e. “Standardized
testing is helping schools improve,” and “Standardized tests help to clarify which learning goals
are most important”). The complete questionnaire can be found in Appendix A.

Constructs

In connection with the results of the literature review, a survey was found that
investigated important areas of testing impact and adapted for this study. The survey questions
centered on four different areas: teachers’ attitudes regarding testing, instructional time, pressure,
and school support. To determine teachers’ attitudes towards testing, the helpfulness construct
questions asked about teachers’ sense of how helpful the testing is in their teaching and
curricular planning ($\alpha=.708$).

Teachers were also asked about how testing impacts their instructional time, including
the influence of testing on teachers’ instructional planning and the class time spent on test
preparation. Earlier studies have shown that teachers often find themselves “teaching to the test,” adjusting their curriculum in order to fit the needs of the standardized assessments. The aim of the class time spent on test preparation construct is to assess to what extent teachers feel that the tests impact their actual classroom instructional focus (α=.810). The influence of testing on teachers’ instructional planning construct measures the impact of the tests on their instructional planning time (α=.877).

Pressure was examined from a variety of sources, including administrators, colleagues, and the community. Due to the most recent policy efforts to attach even more high-stakes accountability sanctions to student test performance, this construct was used to investigate to what extent teachers are feeling this pressure and from whom. The pressure construct measured the extent to which teachers feel pressure from these various sources (α=.805).

Finally, school support was measured by the extent of instructional renewal and school attention to test scores. Given the presence of high-stakes standardized assessments, many studies have identified the importance of a strong administration supporting and collaborating with teachers to develop appropriate educational strategies that benefit students. This can be a challenging task, and so this construct was used to ascertain the success of these strategies as felt by the teachers surveyed. The extent of instructional renewal construct examined how much teachers feel their school administration encourages innovative improvement strategies (α=.718), while the school attention to test scores construct investigates how important teachers believe the test scores are to their administration when considering planning and evaluation (α=.808). These four general factors and the six more specific areas within give an overall sense of the experience of teachers when dealing with high-stakes standardized assessments in their classrooms.
Data Analysis

To analyze the data, an Analysis of Variance (ANOVA), which is a statistical technique for analyzing multiple levels of the independent variable, was used. An ANOVA allows for investigation of similarities and differences between multiple groups from the same population and helps determine whether or not observed differences might have occurred by chance or if they are statistically significant. For the purposes of this study, a one-way analysis of variance was utilized with the level of significance set at $\alpha < 0.10$. Constructs with significant differences amongst groups were identified through a comparison of means.

The independent variables examined included gender, experience, training, subjects taught, and work with gifted and talented students. These were then used to compare the different responses of the dependent variables regarding each of the scales, including testing helpfulness, pride, class time spent on test preparation, influence of testing on teachers’ instructional planning, pressure, extent of instructional renewal, and school attention to test scores.
Chapter 4
Results

This study examined the impact of high-stakes standardized assessments on teachers and their instruction. Overall, the teachers responded that they felt pressure due to the tests and that they impact their instruction, although they did not find the tests particularly helpful in shaping instruction. The responses were also examined for possible differences between the various subgroups of gender, training, experience, subject area, and work with gifted and talented students, finding some differences in gender, training, experience, and student population. While there were several areas of significant difference found between the respondents regarding each of the constructs, also of note were many areas in which differences might be expected but were not found.

Overall Findings

Overall, the teachers responded that they felt a significant amount of pressure surrounding the standardized assessment practices (M=4.33). An overwhelming majority of teachers responded that they strongly agreed that they felt pressure from a variety of sources, including administration, parents, and the community, to improve student test scores. The group also indicated that they felt that the testing significantly impacted their classroom instruction (M=4.02). The respondents also noted that the school was quite focused on test scores (M=3.9) and that they spent a significant amount of class time on test preparation (M=3.87). The teachers also reported feeling a sense of pride in their teaching (M=3.82) and that the school provided a strong sense of instructional renewal (M=3.77). The teachers also indicated that they did not find the tests to be of significant helpfulness in their classroom endeavors (M=2.24).
Group Comparisons

Gender.

Although teaching is a profession typically dominated by females, the school population and sample surveyed allows for some comparison between the two genders given the higher numbers of male teachers represented in both groups. A significant difference (p=.048, F(1,14)=4.778, d=.56) was found amongst the male teachers regarding the class time they spent on test preparation (M=4.44) as compared to female teachers (M=3.58). Thus, male teachers were more inclined to gear their instruction towards the concepts tested on the state-mandated assessments. Further, the male teachers exhibited a somewhat less significant (p=.080, F(1,14)=3.613, d=.52), although still relevant, inclination to focus their instructional planning towards testing (M=4.55) than did female teachers (M=3.75). While both groups indicated a fair amount of testing influence over both planning and classroom instructional time, male teachers expressed a higher likelihood of targeting these activities towards the testing requirements.

Table 3: Gender Comparison

<table>
<thead>
<tr>
<th>Gender</th>
<th>Sample</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Class Time Spent on Test Preparation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Males</td>
<td>4.44</td>
<td>.33</td>
<td>.56</td>
</tr>
<tr>
<td>Females</td>
<td>3.58</td>
<td>.84</td>
<td></td>
</tr>
<tr>
<td>Influence of Testing on Instructional Planning</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Males</td>
<td>4.55</td>
<td>.21</td>
<td>.52</td>
</tr>
<tr>
<td>Females</td>
<td>3.75</td>
<td>.91</td>
<td></td>
</tr>
</tbody>
</table>
Teacher Training/Background.

Respondents were categorized into those who possess a bachelor’s degree and those who possess some additional advanced degree. None of the respondents had earned Educational Specialist or doctoral degrees. Teacher educational background and teacher training are frequently discussed within the discourse surrounding education and reform⁴, and a difference in their responses might be expected. One area of significant difference (p=0.065, F(1,14)=4.068, d=.45) was found in extent of instructional renewal. Teachers with advanced degrees found the extent of instructional renewal present in schools to be slightly less present (M=3.45) than did teachers with a bachelor’s degree only. Despite the differences in their educational training, there were no other significant differences found in their responses.

Table 4: Training Comparison

<table>
<thead>
<tr>
<th>Training/Background</th>
<th>Sample</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Extent of Instructional Renewal</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bachelor’s Degree</td>
<td></td>
<td>4.05</td>
<td>.37</td>
</tr>
<tr>
<td>Advanced Degree</td>
<td></td>
<td>3.45</td>
<td>.75</td>
</tr>
</tbody>
</table>

Experience in Teaching.

There was a wide variety of teaching experience found in the respondents, with a minimum of one year of teaching, including the current school year, and a maximum of 36 years. This range was divided into two subgroups, based on the timeframe of the implementation of NCLB legislation. Groups were defined by those who had taught 10 years or fewer and those

⁴ See Linda Darling-Hammond’s *The flat world and education: How America’s commitment to equity will determine our future* for a thorough discussion on teacher training and education.
who have been teaching for more than 10 years. The two groups demonstrated somewhat little variability in their responses, demonstrating that novice and veteran teachers appear to experience testing in a similar manner, and one group does not have significantly different responses to the implementation of high-stakes tests. A significant difference was found between the respondents within the construct of pride (p=0.084, F(1,14)=3.497, d=0.46).

Teachers who have been teaching for more than 10 years responded that they have a lower sense of pride than teachers who have been teacher for 10 years or less. The veteran group would have been teaching prior to and during the initial implementation of NCLB legislation and mandates.

Table 5: Experience Comparison

<table>
<thead>
<tr>
<th>Experience</th>
<th>Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
</tr>
<tr>
<td>Pride</td>
<td></td>
</tr>
<tr>
<td>10 Years Teaching or Less</td>
<td>4.11</td>
</tr>
<tr>
<td>More than 10 Years Teaching</td>
<td>3.39</td>
</tr>
</tbody>
</table>

Subject Area Focus.

The comparison between teachers who typically teach subjects tested by the state-mandated assessments and those who do not was also examined. Differences might be expected between teachers who test subjects measured by high-stakes standardized assessments and those who do not. Despite experiencing potentially different pressures towards performance given their differing subject areas, teachers who focus on tested subjects did not report any significant differences in their experiences than did those whose subject areas are not tested by state-mandated tests. In each of the construct areas, their responses were quite similar. This result
could indicate that both groups receive equivalent pressure and have similar attitudes towards the testing.

**Gifted and Talented Student Work.**

As the school sampled includes a gifted magnet program, it can also be beneficial to investigate the potential difference in responses between those who work primarily with gifted/talented students versus those whose students typically do not fall into this category. Teachers working with a primarily gifted student population demonstrated a significant difference in responses from those working with the general student population in three different areas: test helpfulness, influence of testing on instructional planning, and extent of instructional renewal.

There was a significant difference in how helpful the two groups of teachers found the tests (p=.009, F(1,14)=9.368, d=.70). Teachers whose students were primarily gifted reported finding the test results less helpful (M=1.6) than did teachers whose students were not classified as such (M=2.57). These results may be expected, as gifted students would be expected to perform well on tests and the test scores may not be as useful for their teachers working to shape curriculum and instruction. There was a somewhat less significant, although still relevance, in how the teachers spent their class time in regards test preparation (p=.061, F(1,14)=4.203, d=.37). Teachers working with gifted students reported that they spent much less class time preparing for the tests (M=3.32) as compared with the other teachers (M=4.14). As their students are likely traditionally more successful on the standardized assessments, the teachers working with primarily gifted students would be expected to tailor instruction towards test practices much less than teachers working with the general student population. Finally, there
was a somewhat significant difference in how the two groups viewed the extent of instructional renewal provided by the school (p=.076, F(1,14)=3.716, d=.43). The gifted teachers reported that they felt the school provided slightly less encouragement for implementing innovative practices (M=3.36) than did the general population teachers (M=3.97). The school may not focus on instructional renewal towards students already successful on the assessments, as one would imagine most gifted students would be.

Table 6: Gifted and Talented Student Work Comparison

<table>
<thead>
<tr>
<th>Gifted and Talented Student Work</th>
<th>Sample</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Test Helpfulness</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teachers with less than 50% gifted students</td>
<td>2.57</td>
<td>.69</td>
<td>.70</td>
</tr>
<tr>
<td>Teachers with at least 50% gifted students</td>
<td>1.60</td>
<td>.15</td>
<td></td>
</tr>
<tr>
<td>Influence of Testing on Instructional Planning</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teachers with less than 50% gifted students</td>
<td>4.25</td>
<td>.61</td>
<td>.37</td>
</tr>
<tr>
<td>Teachers with at least 50% gifted students</td>
<td>3.55</td>
<td>1.10</td>
<td></td>
</tr>
<tr>
<td>Extent of Instructional Renewal</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teachers with less than 50% gifted students</td>
<td>3.97</td>
<td>.46</td>
<td>.43</td>
</tr>
<tr>
<td>Teachers with at least 50% gifted students</td>
<td>3.36</td>
<td>.79</td>
<td></td>
</tr>
</tbody>
</table>
Chapter 5
Conclusion

The results of this study show that teachers are clearly feeling the impact of high-stakes standardized assessments on their teaching. The testing mandates and their punitive sanctions create a pressure-filled environment that impacts teachers’ curriculum and instructional time. Still, though, teachers exhibit a sense of pride in their work, and appreciate opportunities for developing and implementing innovative instructional techniques. These results confirm much of what previous research has demonstrated regarding teachers’ perspectives on high-stakes standardized assessments. However, differences were found amongst subgroup comparisons in gender and work with gifted and talented students that have not previously been addressed in the literature. Further, perhaps even more striking are the areas in which more significant differences might have been expected, but none were found. Differences in teachers’ education levels, experience, and subject areas did not appear to cause significantly different responses, although these characteristics are often debated in the contemporary school reform discourse. These differences and similarities should be examined further, and their impact on teachers and their instruction should be studied. Understanding teachers’ experiences with the tests is important in assisting administrators and policy makers in structuring the policy and support plans that will be most effective.

Overall Results

This survey demonstrates that teachers are feeling a significant amount of pressure from a variety of sources, including their school administration, parents, and the media. The high-stakes testing environment is putting pressure on teachers to ensure their students are successful on the
state-mandated standardized assessments. The survey also shows that these tests are impacting their instruction, in that they have a significant influence on both teachers’ instructional planning as well as actual class time devoted to test preparation activities. These teachers also reported that their school was quite focused on test scores, but also that it placed some emphasis on innovative instructional practices. While the teachers did not report finding the tests helpful in their teaching, they did still exhibit pride in their teaching. The overall results support much of what previous research has found, but the subgroup comparisons lead to several surprising findings. In comparing various subgroup responses, this study found that differences in gender, training, experience, and work with gifted and talented students led to significant differences in responses.

In reexamining the research questions posed, this study has found that teachers typically do not find high-stakes standardized assessments to be helpful in their teaching, especially not teachers working with gifted students. The tests do have a strong impact and influence teachers’ instructional time and classroom practices, especially for male teachers. The teachers also report feeling a great deal of pressure regarding their students’ test scores and improvement from a variety of sources. The teachers in this survey felt that their school administration was quite focused on test scores and was fairly encouraging of innovative instructional practices, but some felt they could be more accepting of these new ideas, in particular, teachers working with gifted and talented students.

As the study demonstrated that teachers feel a tremendous amount of pressure due to high-stakes standardized assessments, district and school administrators should acknowledge the pressure their teachers are feeling and develop strategies to assist teachers with this sense of pressure. The pressure of the tests often negatively impacts teacher performance and motivation,
and teacher motivation is important to student success (Kocabas, 2009; Martin, 2006; Skaalvik & Skaalvik, 2011). The teachers surveyed still demonstrated pride in their work and school, despite the extreme pressure, which the administration can use in building a supportive and collaborative environment. As school leaders are instrumental in creating an atmosphere conducive to learning and producing strong results on assessments (DeMoss, 2002; Reitzug et al., 2008; Singh & Al-Fadhli, 2011; Sunderman et al., 2006), they should be proactive in identifying strategies that will help their teachers cope with the pressures of high-stakes testing and maintain their sense of pride.

Further, school leaders should take an active role in collaborating with teachers to create effective test preparation curriculum. Teachers feel very strongly that testing is impacting their instructional methods and class time. Teachers who have positive feelings about their test preparation programs are typically involved in collaborative settings that develop structured plans for addressing student test preparation needs. This atmosphere then allows them more flexibility and freedom in their other teaching, as less time is devoted strictly to test preparation (DeMoss, 2002; Reitzug et al., 2008; Singh & Al-Fadhli, 2011; Sunderman et al., 2006). As the teachers surveyed voiced strong opinions about the emphasis on the test in their instructional activities, working with them to address those concerns could both reduce pressure and also raise test scores, always still a necessary goal.

Finally, the teachers did not find the assessments to be helpful in their teaching. While some frustration with the tests is to be expected, they still can yield valuable information if the test data is well understood and applied by the teachers. If teachers are educated on what the test data really means, what the tests are meant to measure, and how the information can guide their instruction in meaningful ways, they can find more value in this mandated testing requirement.
While high-stakes standardized assessments provide only one form of assessment, they do provide a measure of student learning that, when combined with additional forms of teacher assessment, can assist teachers in working most effectively with their students. School administrators should ensure teachers have opportunities to be educated about the meaning of test data so that they might benefit from the information it can provide.

**Gender.**

Male teachers reported spending more time on test preparation than did female teachers, and also focused more on testing in their instructional planning than did females. Thus, the male teacher respondents reported the testing more directly impacting their instructional planning and class time than did female teachers. Perhaps male teachers are more likely to emphasize the tests, or find test preparation beneficial to their instruction. Female teachers may see other educational areas as more pressing. The reasons for these differences are unknown and gender differences have not typically been studied. Thus, this will be an important area for future research.

**Training.**

Teachers with bachelor’s degrees responded that they had a greater sense of instructional renewal focus by the school administration than did teachers with further education. Teachers who have pursued education beyond the bachelor’s degree might have developed more innovative strategies they wish to implement, but find the administration less receptive than they might hope, but the true cause of the difference cannot fully be assessed from this survey.
Teacher education might seem likely to lead to significantly different experiences of the testing, as training is frequently the topic of much debate within the discourse (Darling-Hammond, 2010). However, these teachers did not demonstrate many differences based on their educational background. Especially as training and certification options and requirements are being debated, more research should investigate teacher education programs at both the undergraduate and graduate levels to determine what type of instruction will be most beneficial for new teachers entering the contemporary culture of testing. This might indicate an area of concern for graduate schools of education in particular, which may not adequately be preparing students to work with the tests beyond the training they receive in undergraduate studies. For teachers who study areas other than education, specific training and education will also be necessary as they enter the teaching profession, as they will lack the instruction others receive from education-specific classes.

**Experience.**

Veteran teachers responded in similar ways to novice teachers in all areas except pride, with novice teachers indicating stronger feelings of pride. This might be at least partly because the veteran teachers have a different view of the changes brought about by NCLB and the discourse of failure, as they taught prior to and during its implementation. Still, greater differences in other constructs might have been expected. Little research has focused on examining teaching experience and testing. However, it is commonly discussed in school reform debates as novice teachers are often criticized for a lack of experience but also praised for the promise of energy and fresh new ideas. Veteran teachers are both hailed for their commitment and experience, and also criticized for being stuck in their ways or only teaching due to tenure.
Perhaps these two groups are actually quite similar, and the assumptions of the discourse are flawed. This will be an important area for further research.

**Gifted and Talented Work.**

This study also found significant differences in the responses between teachers who work primarily with gifted and talented students, and those who work primarily with the general student population. Specifically, teachers of gifted students found the tests less helpful and less influential on their instructional planning, and also reported finding that the school was less supportive of innovative teaching practices.

The causes of the differences in responses cannot be assessed from this study. These results may indicate that the gifted and talented students are typically successful on the tests and their scores are not helpful due to their exceptionalities. Their instructors may be particularly interested in innovative teaching strategies in an effort to meet the students’ special needs and are therefore more sensitive to the administrative support they receive. Whatever the causes, research should be done in this area. Oftentimes students who struggle on these tests receive more attention in research, but it is important to examine the impact of the tests on this group of students and their teachers as well.

The school studied and others like it who also have a magnet program within the school should work to make sure they take advantage of the magnet program for the benefit of all students. The teachers surveyed who work largely with gifted students displayed feelings of less encouragement for innovative instructional practices than did their counterparts working with the general school population. Encouraging these teachers to be more innovative, especially paired with an emphasis on collaboration, could work to exploit the strengths of the magnet program.
and improve learning for all students. Administrators should work to create opportunities for collaboration amongst teachers that allow teachers working with gifted and talented students to implement innovative instructional techniques alongside teachers working with other students so that both can feel encouraged and benefit from new and creative strategies.

**Discussion**

Testing has enjoyed a steady rise in influence over the past century in education in America due to a variety of factors, including the learning sciences and behaviorism, the feminization of the teaching profession, the eugenics movement, military concerns, an interest in equity, and a focus on global competition. As we trace the evolution of testing and its progression into what is now an integral and assumed part of public education in the United States, we can imagine that it will remain influential in our education system for some time.

While many undertake efforts to dissect the validity of the tests, the achievement gap (Arbuthnot, 2011; Taubman, 2009), and other concerns regarding accountability measures based on these tests, we must also address the very real consequences of the tests as they impact teachers in the classroom. Efforts to challenge the testing system are certainly underway and will continue, but as those forces work to change the system, additional efforts must be made to assist teachers working within the current state of affairs.

Early in the twentieth century, as public schooling in America was rapidly expanding, a scientific interest in people regarding psychology and social engineering was also developing. Researchers created new tests and mechanized systems of learning. Even though experimental school-based curriculum was found most successful in preparing students for success in college (Lagemann, 2000), standardization and testing became dominant in education. Though these
very tests have been found to highlight achievement gaps for almost 50 years, particularly surrounding socioeconomic status, these findings seem less important in contemporary discourse, as teachers are instead the focus regarding student achievement. While concerns regarding student achievement persist, the debate surrounding education and reform emphasizes accountability, standardization, and data-driven assessment. Teachers are blamed for low student scores, while out-of-school contextual factors impacting student performance are essentially ignored.

Standards and accountability now dominate educational discourse, while teachers are often marginalized. Even while teachers protest on the steps of capitol buildings across the country, legislatures continue to enact policies that run counter to teachers’ concerns. Discourse surrounding teachers is often quite harsh, criticizing them as both lazy and greedy. They are lambasted as being against any form of evaluation or accountability in education, which, as the literature shows, is clearly untrue. These problematic generalizations discount and diminish the valid arguments teachers actually make about each of these issues, as policy makers move forward on the public current demanding higher standards and better scores. Still, teachers work throughout the turmoil and return to the classroom day after day.

Thus, testing continues, and teachers continue to work within a system they often see as problematic. Continuing to learn about teachers’ experiences of high-stakes standardized assessments in their classrooms, as well as the perspectives of students, parents, and administrators, will be important for determining the needs of the most central participants in the education system. While one can hope that education can evolve its focus on assessment to be broader than just standardized assessments when considering accountability and high-stakes consequences for schools, proper support needs to be available for teachers currently working in
the public schools. These teachers must deal with the tests, and so any efforts to make the testing process and its results more meaningful and helpful for teachers can benefit them and their instruction.

Schools can take advantage of the mandates by developing support programs for teachers that assist them in understanding and utilizing test data to refine existing instructional methods and to create and implement new and innovative strategies as well. While these test scores are only one measure of student learning, they do provide information teachers can utilize if they are given the proper information about what the tests mean and then given the opportunity to collaborate and innovate. Teachers are proud of their profession and excited about working with students, and should be given the opportunity to engage with each other and the curriculum. While teachers can often be burdened by the demands of the job and the bureaucracy surrounding it, their administrators should work to alleviate those burdens in order to allow the space for engagement and collaboration. Educators should treat the data provided by test scores as just one more assessment tool within the larger framework of curriculum and curriculum development within the school.

School administrators might consider administering similar surveys to that used in this study in order to ascertain the particular opinions of the teachers in their schools. This could help them understand how their staff perceives the impact of high-stakes standardized assessments at their school, which can help target specific changes to be made, or programs to be implemented, in order to address teacher concerns. This can help them assess not only particular concerns teachers have, but also specific strengths on which they can capitalize, such as magnet programs, teachers with particular types of experience, or community factors. Further,
administrators must be open to the results, and be ready to engage teachers fully in any plans moving forward.

Future research should continue to uncover the consequences of the testing movement in order to shape the best educational supports and policies. Further, we must continue to question the test itself, especially as certain types of testing are used to the exclusion of all other types of assessments for accountability purposes. Insights regarding teachers’ experiences of the reform policies put in place can be useful in providing the tools and supports they need, and also to evaluate the system as a whole.

Limitations

There are several limitations regarding the generalizability of this study. The survey sample was quite small, as only 15 people responded. Thus, their responses may not be applicable in a larger population. Further, the respondents were all from only one school. This is a limitation because they are influenced by their school administration and the contextual factors surrounding them, such as the parents and students with whom they work, as well as a very particular community culture which may lead to results different from those of another group.

Because the teachers were asked to voluntarily respond to an online survey, those who are less comfortable using technology, or those who have concerns regarding the confidentiality of online surveys, may not have chosen to participate. Further, the survey was administered within the few weeks leading up to the annual state-mandated assessments, which may have impacted responses or limited respondents who were overwhelmed with the requirements of
administering the assessments. Still, though, learning about the perspectives of this, or any, teacher group can be beneficial in understanding the true impact of educational reform policies.
References


Appendix A: Survey

The purpose of this survey is to examine teachers’ opinions of the impact of standardized tests on their schools and teaching. This survey should take no longer than 20 minutes. Your participation in this survey is greatly appreciated.

1. What is your gender?
   Male   Female   Transgender

2. What is your age?

3. What is your Race/Ethnicity? Please select all that apply:
   American Indian   Native Hawaiian or Other Pacific
   Asian   Islander
   Black or African American   White or Caucasian
   Hispanic   Other (please specify): ____________

4. Grade(s) you currently teach:
   6   7   8

5. Grade(s) you taught last year:
   6   7   8

6. Years you've been teaching (including this year): _________years

7. Years you've been teaching at this school (including this year): ______ years

8. Education/Training (please check all that apply):
   Four-year college degree/BA/BS in education
   Four-year college degree/BA/BS in area other than education
MS/MA in education

MS/MA in area other than education

Education Specialist

Doctorate

Alternative Certification

Other Education/Training (please specify): ____________

9. What subject area(s) do you teach?

   English/Language Arts       Foreign Languages
   Math                      Arts
   Science                    Other (please specify): ____________
   Social Studies

10. How many of your students are gifted/talented/mastery students?

   0-25%       51-75%
   26-50%      76-100%

Please respond to each of the statements below, indicating the extent to which you agree or disagree with each one using the following scale:

   1 — Strongly Disagree
   2 — Disagree
   3 — Neutral
   4 — Agree
   5 — Strongly Agree
11. My current students' academic ability level is comparable to the national norm.
12. My current students' academic ability level is comparable to students I have had over the last three years at the same grade level.
13. I expect my students to perform well on standardized tests.
14. Many of the students I teach are not capable of learning the material I am supposed to teach.
15. Teachers can influence substantially how well their students do on standardized tests.
16. Standardized testing is helping schools improve.
17. Standardized testing creates a lot of tension for teachers.
18. Standardized testing creates a lot of tension for students.
19. Standardized tests give me important feedback about how well I am teaching in each curricular area.
20. Standardized tests help to clarify which learning goals are most important.
21. Staff feel there is a discrepancy between what they think should be taught and what the standardized tests emphasize.
22. I sometimes feel it is a waste of time to try to do my best as a teacher.
23. Standardized test results give an accurate reading of student learning.
24. Most of our school staff members have a strong sense of pride in their work.
25. Our school is more interested in increasing standardized test scores than in improving overall student learning.
26. I spend time giving students worksheets that review expected standardized test content.
27. I spend time giving students practice in the kinds of item formats that are on the 
standardized test.
28. I spend time giving students commercially-produced practice standardized tests.
29. I spend time giving students old forms of the standardized test on which to practice.
30. I spend time instructing students on standardized test-taking strategies.
31. I look at old or current standardized tests to make sure that my curriculum includes all or 
most of the standardized test's content.
32. I make sure the objectives of the standardized test are covered in my instruction.
33. I adjust my instructional plans based on the standardized test performance of the class I had 
last year.
34. I adjust my instructional plans based on my current students' most recent standardized test 
results.
35. I adjust the sequence of my curriculum based on what's included in the standardized test.
36. I give attention to higher-order thinking and problem-solving skills in my classroom.
37. I drill students in basic skills (e.g., vocabulary, grammar, computations) in my classroom.
38. I give attention to fine arts (music, art) in my classroom.
39. I give attention to science in my classroom.
40. I give attention to subjects which are not standardized tested in my classroom.
41. I give attention to standardized test preparation (homework and class work) in my 
classroom.
42. I feel pressure from my principal to improve my students’ standardized test scores.
43. I feel pressure from school administrators other than the principal to improve my students' standardized test scores.

44. I feel pressure from other teachers to improve my students' standardized test scores.

45. I feel pressure from the district administration/board to improve my students' standardized test scores.

46. I feel pressure from parents to improve my students' standardized test scores.

47. I feel pressure from the community to improve my students' standardized test scores.

48. I feel pressure from the newspaper/media to improve my students' standardized test scores.

49. I feel pressure from my students to improve their standardized test scores.

50. Programmatic efforts to improve student learning are present in my school.

51. Implementation of innovative instructional strategies is present in my school.

52. Support for school-wide or grade-level planning is present in my school.

53. School or grade-wide efforts to improve school or class climate are present in my school.

54. Students in my school are interested in learning.

55. Opportunities for students to choose what they want to study are present in my school.

56. Students' pride in school is present in my school.

57. My school lets teachers know how their students performed compared to other teachers.

58. My school considers standardized test scores to be very important when evaluating teachers.

59. My school administration holds staff meetings to review standardized test scores.

60. My school administration discusses ways to improve standardized test scores.
61. My school administration discusses ways to strengthen instruction in the specific areas where standardized test scores show weakness.

62. My school administration provides materials to give students practice in standardized test-taking skills.

63. My school administration provides special assistance to help individual teachers improve standardized test scores.

64. My school administration checks to see that teachers are emphasizing skills which showed weakness from past standardized test results.

This survey or survey items are used with permission from The National Center for Research on Evaluation, Standards, and Student Testing (CRESST). Copyright © 1991 and by The Regents of the University of California as supported under the Institute of Education Sciences (IES), U.S. Department of Education.
Vita

Jennifer Timmer is currently pursuing a master’s degree in curriculum theory from Louisiana State University. She is also the Associate Director of Student Learning at the Academic Center for Student-Athletes at Louisiana State University. She joined the ACSA in August of 2005 as a Learning Specialist and in August of 2008, she was named the Director of Directed Studies. She earned a master’s degree in English from the University of Miami and a bachelor’s degree in English and math from the University of Kansas.