AN HISTORICAL AND LEGAL REVIEW OF HIGH-STAKES TESTING

By

Michael J. Krolak

(Under the Direction of John Dayton)

ABSTRACT

The high-stakes testing of students is now a multi-billion dollar industry with multi-billion dollar consequences. Since the provisions of the NCLB ratcheted up accountability for students, educational stakeholders and observers have railed against the current high-stakes testing model. High-stakes testing is an industry whose affects are felt worldwide, as not only students and teachers are compared across districts and states, but whole countries are compared. Educational stakeholders across the globe wait with anticipation to see how they rank. Many argue against the standardized tests and the consequences of the results.

High-stakes testing is, "When significant educational paths or choices of an individual are directly affected by test performance, such as whether a student is promoted or retained at a grade level, graduated, or admitted or placed into a desired program, the test use is said to have high-stakes" (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 1999, p. 139). These tests are usually given on a statewide basis under strict security measures. The high-stakes are for the student and the school. The student results of the test can determine whether the student is promoted to the next

grade level or if they have earned a high school diploma. The schools are to measure how well the school's students compare against the other schools and against state standards.

This study provides an historical and legal review of high-stakes testing and how educational testing has transformed into a pass/fail anxiety-inducing test for the current generation of students. High-stakes testing did not just appear overnight. There are landmark cases and state and federal legislation that allowed high-stakes testing to develop. The purpose of this study was to review important historical and legal events related to high stakes testing and its impacts on students.

INDEX WORDS: High-stakes testing, federal legislation, standards, fairness, legal analysis

AN HISTORICAL AND LEGAL REVIEW OF HIGH-STAKES TESTING

By

MICHAEL J. KROLAK

B.B.A., Florida Atlantic University, 1986

M. Ed., University of Georgia, 2009

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF EDUCATION

ATHENS, GEORGIA

2013

AN HISTORICAL AND LEGAL REVIEW OF HIGH-STAKES TESTING

By

MICHAEL J. KROLAK

Major Professor:	John Dayton
Committee:	Elizabeth DeBray
	Sheneka Williams

Electronic Version Approved:

Maureen Grasso

Dean of the Graduate School

The University of Georgia

December 2013

DEDICATION

To the greatest, strongest, funniest, honest person I know my Mom, Marge Krolak. For without her, I would not be here to write these words.

ACKNOWLEDGEMENTS

I would like to acknowledge all the people that have helped me pursue my educational dreams. First and foremost are my parents Marge and Joe Krolak, your love over the years has inspired me to be the best person I can be, myself. I only hope that I have measured up to your expectations. Dad, I miss you and my only regret is that you were not here to share in this process of making you proud. To my sister, Peggy you are the best sister in the world, I hope you realize that and I love you.

I also would like to acknowledge my teachers, professors, friends, students and colleagues along the way that have endured my antics yet continued to support my efforts. To my editor, Donna Geisinger, a million thanks for the million words of advice and edits. Thank you to my committee of Dr. John Dayton, Dr. Elizabeth DeBray and Dr. Sheneka Williams for your belief, trust and encouragement.

Lastly, I would like to acknowledge my two most important, favorite people in the world, my wife Julie and my daughter Amber. Amber, you have shown me how perseverance, fortitude along with beauty inside and outside goes a long way. I only hope you know how much I love you and Jaiden. I am proud of you Amber to infinity and beyond.

Julie, you are my rock and my friend. You are the reason I never quit nor gave up. I will be forever indebted to you for your undying love and patience during this high-stakes, high-anxiety inducing endeavor. Without you by my side, all this hard work would be meaningless. I am a better person for knowing you. Thank you and I love you. We have truly just begun.

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER 1

INTRODUCTION

Statement of the Problem

The high-stakes testing of students is now a multi-billion dollar industry with multi-billion dollar consequences. Since the provisions of the NCLB ratcheted up the accountability of students, educational stakeholders and observers have railed against the current high-stakes testing model. In 2001 in Massachusetts a group of union and educational leaders called New Democracy suggested collective non-compliance: “They propose that teacher union locals vote to refuse to administer the test as a union action—analogous to a work stoppage” (Stratman, 2001, p. 56). Just recently Dornfield (2013) noted, “Teachers at a school in Seattle have had enough. They're refusing to give their students a standardized test that's required by the district. The reason: The test is useless, they say, and wastes valuable time” (Dornfield, 2013, p. 1). But for-profit high-stakes testing will be difficult to defeat. “100 million standardized tests are administered to students each year” (Medina & Neill, 1990, p. 1) at a cost of “$200-500 billion dollars a year” (Baines & Stanley, 2004, p. 8) No consensus on a solution exists. And many are unaware of the legal and political framework in place that has allowed high-stakes testing to evolve.

High-stakes testing is commonly viewed as a relatively new educational reform term. In an online interview of David Koretz, Koretz stated: “It's new in its extremity and its current forms, but it dates back at least twenty-five years or more. There are some tests that go back further; the New York States Regents tests go back 140 years” (Koretz, 2002). What is new is

how high-stakes testing has changed the educational landscape by attaching so much weight to the high-stakes test itself. “Stakes are a powerful lever for affecting change, but one whose effects is uncertain; and that one-size-fit all model of standards, tests, and accountability is unlikely to bring about the greatest motivation and learning for all students (Clarke et al., 2003, p. 5).

The additional weight placed on high stakes testing has escalated questions of fairness from concerns to legal challenges. Most issues of fairness stem from the public’s perceptions of the high-stakes test. People have a myriad of questions such as;

> If the nation’s goal is a high-quality education for all, why not use assessments that can at least tell us if that goal is being met? Why not rely on multiple sources of evidence to inhibit narrowing curriculum and teaching to one test format? Why not make decisions about students and schools based on information gathered over time? Why not transform assessment and accountability to serve the educational needs of all students? (O’Neill, 2003, p. 28).

Koretz also stated in his 2002 interview “How can we include tests in accountability systems in ways that minimize some of these undesirable effects and maximize what we gain?” (Koretz, 2002). “Are these tests a fair measure of student achievement? Is it fair to allow a failing score on a single exam to trump years of good grades? Does the test discriminate against African-Americans or English Language Learners or students with disabilities?” (O'Neill, 2003, p. 634).

High-stakes testing is an industry whose affects are felt worldwide, as not only students and teachers are compared across districts and states, but whole countries are compared. Educational stakeholders across the globe wait with anticipation to see how they rank. Many

argue against the standardized tests and the consequences of the results. High-stakes testing is, "When significant educational paths or choices of an individual are directly affected by test performance, such as whether a student is promoted or retained at a grade level, graduated, or admitted or placed into a desired program, the test use is said to have high-stakes" (American Educational Research Association et al., 1999, p. 139). These tests are usually given on a statewide basis under strict security measures. The high-stakes are for the student and the school. The student results of the test can determine whether the student is promoted to the next grade level or if they have earned a high school diploma. The schools are to measure how well the school's students compare against the other schools and against state standards. The schools are also measured against national standards including those set by the No Child Left Behind Act of 2001(NCLB). Under the NCLB the achievement of standards results is not just about prestige, but in dozens of states these tests measure out how much funding the schools will receive. The NCLB includes punitive consequences if AYP (Adequate Yearly Progress) is not met.

While many educational stakeholders agree that assessments need to be done to evaluate student learning, high-stakes testing are not always fair or legal. For example, there have been many specific challenges to some aspect of high-stakes testing such as:

> Racial and cultural bias *(Larry P. v. Riles*); linguistic bias in tests (*Diana v. California State Board of Education*); test results that dominate special education placement decisions (*Larry P. v. Riles*); failure to test sufficiently or at the appropriate time (*Hoffman v. Board of Education of New York City*); unequal opportunities to learn tested material (*Debra P. v. Turlington*); assignment to ability tracks without educational justification (*Dillon County School District*); racial discrimination in the interpretation of college admission test results (*Regents of the State of California v. Bakke*); gender

discrimination in the interpretation of test results (*Bray v. Lee*); use of the wrong type of test (*Sharif v. New York State Education Department*); use of test results to deny education (*Pennsylvania Association for Retarded Children v. Commonwealth of Pennsylvania*); and access to student academic records (*Family Education Rights and Privacy Act of 1974*) (Childs, Eric Clearinghouse on Tests, Evaluation, & American Institutes for Research, 1990).

Many of the above cases are brought to courts due to the perception that a specific segment of the student population is being harmed. Some of these cases have been cited as landmark cases for the segment of students affected.

Purpose Statement

The purpose of this study was to review important historical and legal events related to high-stakes testing and its impacts on students.

The study is a legal review of high-stakes and how educational testing was transformed into a pass/fail anxiety-inducing test for the current generation of students. High-stakes testing did not just evolve overnight; there are landmark cases, and state and federal legislation that allowed high-stakes testing to evolve.

Research Questions:

This study addressed the following research questions:

1. What is the relevant legal history of high-stakes testing?
2. What are the predominate issues that have been litigated in the courts relevant to high-stakes testing?
3. What is the current legal status of high-stakes testing?

Research Design

This study reviewed the relevant legislation and court cases affecting high-stakes testing since 1950.

Legal research framework

As Dayton noted, "Legal research methods are by design straightforward and flexible to promote accuracy and efficiency. Overreliance on methodology can cause the researcher to miss important evidence and conclusions. Methodology alone is no substitute for thoroughness, reason, common sense, and academic integrity in research"(Dayton, 2013, p. 5).

"Legal research involves a process of reduction, carefully sorting materials starting with all potentially relevant materials down to confirmed relevant materials for analysis and then synthesis into a coherent, current snap-shot picture of the law"(Dayton, 2013, p. 8).

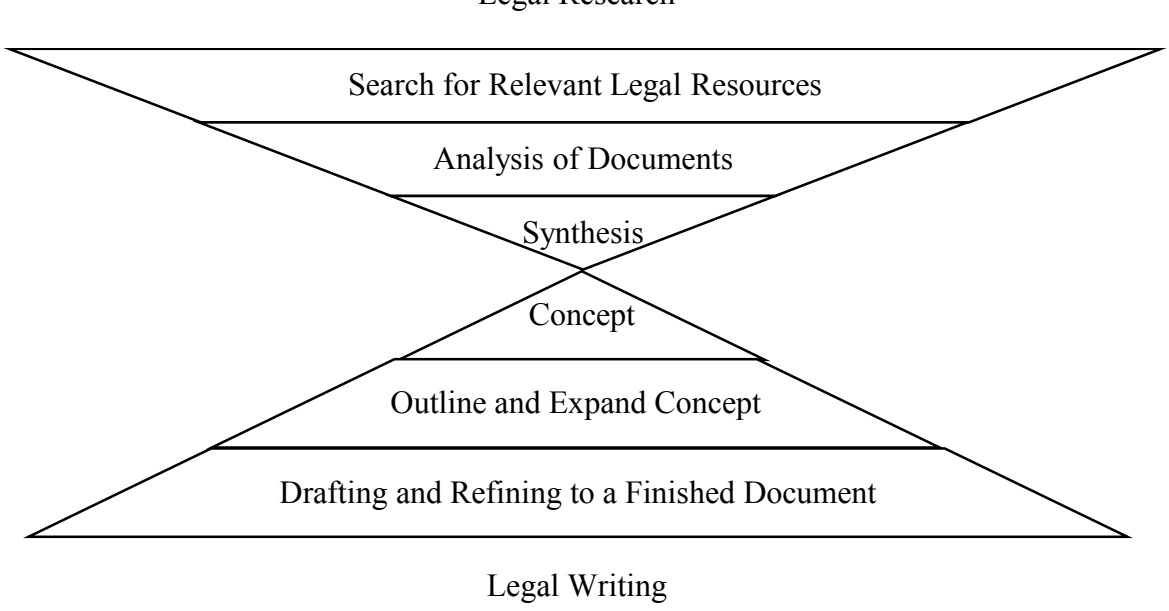

Figure 1.1

The Combined Legal Research and Writing Process according to Dayton

In this study the basic Search for Relevant Resources was done by searching for cases of high-stakes testing being challenged citing equal protection or due process issues as outlined in the U.S. Constitution. The main resource was the cases themselves or the legislation that influenced those cases. Another resource used was law reviews and scholarly reviews of the cases.

The next step was to analyze Relevant Legal Resources. At this phase, the cases or legislation are on point and require in depth analysis and understanding, "After sorting and discarding irrelevant documents, thoroughly review relevant documents"(Dayton, 2013, p. 12).

In the final step, Synthesize means "Synthesize findings from your analysis of documents into a logical, coherent, and concise review of your issue in order to answer your legal question"(Dayton, 2013, p. 12). This was done by explaining the cases and how they are cogent to the study. This study will use the IRAC method employed by Dayton (2013). The IRAC method is a research and writing system for explaining the case; Identify the issue, the Rule of law of the particular case, the Application of law in the case and the Conclusions of the case.

Limitations

The first and immediate limitation of the study is the limit of the scope. Standardized testing has been around for decades and thus has been challenged for decades. This study reviewed legislation and court cases that are broad in scope and address high-stakes testing. There are a multitude of cases that have affected high-stakes testing. Examining every case filed against a high-stakes issue would not be feasible since there are as many cases as there are states, laws and educational segments of students. This study focuses on cases of national importance.

CHAPTER 2

Literature Review

Classroom assessment can yield invaluable information to both the student and teacher. The information gained can be utilized to confirm the teaching and learning done in the classroom. The challenge of high-stakes testing has been to meet these goals. "A 'high-stakes test' is a test that plays a significant role in a significant educational decision. When such a test is used inappropriately, it fails to be beneficial" (Mueller, 2001, p. 202).

High-stakes testing is a complex issue with many facets that has encountered both legal and legislative issues.

> "Untidy fallout from the interaction between litigation and public policy is common in many policy sectors, especially education. With education policy in particular, this untidiness results partly from the inherent complexity of numerous education policies as well as from the importance of the stakes involved" (Heise, 2009, p. 327).

Along with the intricacies of the law, legislators and judges have taken into consideration, at times, the standards a fair assessment is to adhere to. These fairness standards are the frameworks of how an unbiased test should be developed.

This chapter can be divided into two main sections. The first section is a review of the standards of a fair assessment. These standards are the outlines for a fair assessment and are to be the guidelines for a fair high-stakes test. The second section contains the legislation and legal cases that affected high-stakes tests and a review of relevant scholarly commentary. The laws

passed and the court decisions over the last fifty years have helped structure the high-stakes testing environment of today.

Assessment Standards

What should assessment look like?

Before a study can be conducted about an analysis of the laws and legislation enacted, one must first look at the record of what assessments should look like. There are many opinions about assessments but one must look at the scholarly advice of what standards assessments should adhere.

Standards for Educational Research and Testing

There have been five revisions of this report done by the "Joint Committee on Standards for Educational and Psychological Testing of the American Educational Research Association, the America Psychological Association, and the National Council on Measurement in Education" (American Educational Research Association et al., 1999, p. ii). The first report was done in 1954 and was called *Technical Recommendations for Psychological Tests and Diagnostic Techniques*. The subsequent revisions were retitled, *Standards for Educational and Psychological Testing* (*Standards*). As high-stakes testing has evolved so has the *Standards*. The reasoning for this committee to produce these standards are; "To promote the sound and ethical use of tests and to provide a basis for evaluating the quality of testing practices" (American Educational Research Association et al., 1999, p. 1).

When it comes to the general subject of educational testing, one must first define what constitutes high-stake testing. The committee defined high-stakes testing as,

> When significant educational paths or choices of an individual are directly affected by test performance, such as whether a student is promoted or retained at a grade level,

> graduated, or admitted or placed into a desired program, the test use is said to have high-stakes (American Educational Research Association et al., 1999, p. 139).

The committee also recognized some caveats that go along with high-stakes testing,

> The higher the stakes associated with a given test use, the more important it is that test-based inferences are supported with strong evidence of technical quality. The tests needs to exhibit higher standards of technical quality for its avowed purposes that might be expected for lower-stakes purposes. Efforts need to be made to minimize errors in estimating individual scores or in classifying individuals in pass/fail or admit/reject categories (American Educational Research Association et al., 1999, p. 139).

This shows the higher expectations from many facets of the high-stakes community. The higher the standard the higher the responsibilities to make sure tests are designed, administered and assessed properly.

The latest printed version of the *Standards* was in 1999; however there is an updated version which is expected to be printed sometime in 2013. According to the website of the American Psychological Association (APA), an update written by APA members of the committee, does recognize the shortcomings of the jointly written book(s), "There is no mechanism to enforce compliance with the *Standards* on the part of the test developer or test user" (Camara & Ernesto, 2011). While *Standards* are a suggestion, the authors illustrate that *Standards* has been influential,

> However, the *Standards* have been referenced in federal law and cited in Supreme Court and other judicial decisions, lending additional authority to the document. For example, they have been cited in the Goals 2000: Educate America Act and Title I of the Elementary and Secondary Education Act. They were also cited in several major court

decisions involving employment testing, including a Supreme Court case in 1988 (Camara & Ernesto, 2011).

An updated version of the *Standards* on the APA website has the summarized propositions;

> Primary differences between the revised draft of the Standards and the 1999 edition include the separation of the general text into 'foundations,' 'operations,' and 'applications' sections; a condensation of several individual chapters dealing with issues of fairness into a single 'foundations' chapter; the inclusion of topics such as educational accountability and technological advances in testing; and the re-organization of chapters concerning workplace testing and credentialing. The decision to condense the multiple chapters that deal with fairness issues in the 1999 Standards into one 'foundations' chapter in the new edition was made after members of the Joint Committee concluded that the issue of fairness was so fundamental to testing practice, it should be considered as a foundation of testing along with validity and reliability. This change means that the revised Standards will contain an expanded discussion of fairness, and the total number of chapters will decrease from the current total of 15 to 13 (Camara & Ernesto, 2011).

Therefore since the printed *Standards* of 1999 have been agreed upon by the joint committee, this study looked at the most recent printed *Standards* of 1999 instead of proposed updates to the 2013 version which have not been approved nor printed yet. The test standards were laid out in broad categories with the more specific details within those sections;

Part I: Test Construction, Evaluation, and Documentation

Validity

Reliability and Errors of Measurement

Test Development and Revision

Scales, Norms, and Score Comparability

Test Administration, Scoring, and Reporting

Supporting Documentation for Tests

Part II: Fairness in Testing

Fairness in Testing and Test Use

The Rights and Responsibilities of Test Takers

Testing Individuals of Diverse Linguistic Backgrounds

Testing Individuals with Disabilities

Part III: Testing Applications

The Responsibilities of Test Users

Psychological Testing and Assessment

Educational Testing and Assessment

Testing in Employment and Credentialing

Testing in Program Evaluation and Public Policy (American Educational Research Association et al., 1999, p. iv).

Since one of the foci of this dissertation is the fairness of high-stakes testing, the interpretations of the standards were done on the sections that had to do with fairness in testing. The committee reasoning for addressing fairness is stated as, "It is intended to emphasize the importance of fairness in all aspects of testing and assessment and to serve as a context for the technical standards" (American Educational Research Association et al., 1999, p. 73). This joint committee also details their focus on fairness in testing;

> The focus of the *Standards* is on those aspects of tests, testing, and test use that are the customary responsibilities of those who make, use, and interpret test, and that are characterized by some measure of professional and technical consensus (American Educational Research Association et al., 1999, p. 73).

This committee attempted to come up with a concise definition of fairness as it has to do with this type of testing. They opine about the difficulty with having one definition of fairness "It is possible that two individuals may endorse fairness in testing as a desirable social goal, yet reach quite different conclusion about the fairness of a given testing program" (American Educational Research Association et al., 1999, p. 74). They wrote about their attempts to define fairness by breaking their ideas into categories with an explanation of these categories;

Fairness as Lack of Bias

> It is said to arise when deficiencies in a test itself of the manner in which it is used result in different meanings for scores earned by members of different identifiable subgroups…there is general consensus that consideration of bias is critical to sound testing practice.

Fairness as Equitable Treatment in the Testing Process

> There is also consensus that fair treatment of all examinees requires consideration not only of a test itself, but also the context and purpose of testing and the manner in which test scores are used.

Fairness as Equality in Outcomes of Testing

> Given evidence of the validity of intended test uses and interpretations, including evidence of lack of bias and attention to issues of fair treatment, fairness has been established regardless of group-level outcomes.

Fairness as Opportunity to Learn

> When test takers have not had the opportunity to learn the material tested the policy of using their test scores as a basis for withholding a high school diploma, for example, is viewed as unfair (American Educational Research Association et al., 1999, p. 76).

While *Standards* does address fairness and bias in testing the committee does admit they are far from having all the answers, "*Standards* will provide more specific guidelines on matters of technical adequacy, matters of values and public policy are crucial to responsible test use" (American Educational Research Association et al., 1999, p. 80).

The twelve specific fairness standards are listed in *Standards* as:

Standard 7.1

When credible research reports that test scores differ in meaning across examinee subgroups for the type of test in question, then to the extent feasible, the same forms of validity evidence collected for the examinee population as a whole should also be collected for every relevant subgroup. Subgroups may be found to differ with respect to appropriateness of test content, internal structure of test responses, the relation of test scores to other variables, or the response processes employed by individual examinees. Any such finding should receive due consideration in the interpretations and use of scores as well as in subsequent test revisions.

Standard 7.2

When credible research reports differences in the effects of construct-irrelevant variance across subgroups of test takers on performance on some part of the test, the test should be

used if at all only for those subgroups for which evidence indicates that valid inferences can be drawn from test scores.

Standard 7.3

When credible research reports that differential item functioning exist across age, gender, racial/ethnic, cultural, disability and/or linguistic groups in the population of test takers in the content domain measured by the test, test developers should conduct appropriate studies when feasible. Such research should seek to detect and eliminate aspects of test design, content, and format that might bias test scores for particular groups

Standard 7.4

Test developers should strive to identify and eliminate language, symbols, words, phrases, and content that are generally regarded as offensive by members of racial, ethnic gender, or other groups except when judged to be necessary for adequate representation of the domain.

Standard 7.5

In testing applications involving individualized interpretations of test sores other than selection, a test takers' score should not be accepted as a reflection of standing on the characteristic being assessed without consideration of alternate explanations for the test takers' performance on that test at that time.

Standard 7.6

When empirical studies of differential prediction of a criterion for members of different subgroups are conducted, they should include regression equations (or appropriate

equivalent) computed separately for each group of treatment under consideration or an analysis in which the group or treatment variables are entered as moderator variables.

Standard 7.7

In testing applications where the level of linguistic or reading ability is not part of the construct of interest, the linguistic or reading demands of the test should be kept to the minimum necessary for the valid assessment of the intended construct.

Standard 7.8

When scores are disaggregated an publicly reported for groups identified by characteristics such as gender, ethnicity, age language proficiency, or disability, cautionary statements should be included whenever credible research reports that test scores may not have comparable meaning across these different groups.

Standard 7.9

When tests or assessments are proposed for use as instruments of social educational or public policy, the test developers or users proposing the test should fully and accurately inform policymakers of the characteristics of the tests as well as any relevant and credible information that may be available concerning the likely consequences of test use.

Standard 7.10

When the use of a test results in outcomes that affect the life chances or educational opportunities of examinees, evidence of mean test score differences between relevant subgroups of examinees should, where feasible, be examined for subgroups for which credible research reports mean differences for similar tests. Where mean differences are found, an investigation should be undertaken to determine that such differences are not attributable to a source of construct underrepresentation or construct-irrelevant variance.

> While initially the responsibility of the test developer, the test users bears responsibility for uses with groups other than those specified by the developer.
>
> Standard 7.11
>
> When construct can be measured indifferent ways that are approximately equal in their degree of construct representation and freedom from construct-irrelevant variance, evidence of mean score differences across relevant subgroups of examinees should be considered in deciding which test to use.
>
> Standard 7.12
>
> The testing or assessment process should be carried out so that test takers receive comparable and equitable treatment during all phases of the testing or assessment process (American Educational Research Association et al., 1999, pp. 80-84).

This shows a complete set of standards that just address fairness alone that could be used as the template for any test, legislation or legal opinion. While this particular set of standards only address general fairness in testing and testing use, the committee designed additional fairness standards to apply to specific groups titled, "The Rights and Responsibilities of Test Takers, Testing Individuals of Diverse Linguistic Backgrounds and Testing Individuals with Disabilities" (American Educational Research Association et al., 1999, p. iv). The committee members apparently saw the need to address the broad impact of fairness in high-stakes testing in these areas.

Standards set two hundred sixty four educational standards designed to help all test designers, implementers, assessors, takers and users. Sixty seven standards have to do with fairness or educational testing, forty eight standards just for fairness alone. This shows that there is a researched basis to be used in considering most if not all testing questions. The committee

shows its willingness to evolve these standards since they revisit and update the standards as needed.

High Stakes: Testing for Tracking, Promotion, and Graduation

This report was written at the behest of then President Clinton to evaluate testing use. Voluntary National Testing (VNT) was a provision from the Improving America's Schools Act (IASA). "…Congress and the Clinton administration asked the National Research Council, through its Board on Testing and Assessment (BOTA), to conduct three fast-track studies over a 10-month period" (Heubert, Hauser, & National Academy of Sciences - National Research Council, 1999, p. 5). A prestigious group of researchers were to investigate the uses of testing and make recommendations on any changes that should be implemented. The members of this research group were well recognized authorities within their fields;

> The project that is the subject of this report was approved by the Governing Board of the National Research Council, whose members are drawn from the councils of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine (Heubert et al., 1999, p. ii).

It was the intention of the Clinton administration to quell the arguments of high-stakes testing by using this scientific panel to validate what was right and wrong with testing. Specifically, the committee was mandated with the following tasks:

> Public Law 105-78, enacted November 13, 1997
>
> SEC. 309. (a) STUDY—The National Academy of Sciences shall conduct a study and make written recommendations on appropriate methods, practices, and safeguards to ensure that—

> (1) existing and new tests that are used to assess student performance are not used in a discriminatory manner or inappropriately for student promotion, tracking or graduation; and
>
> (2) existing and new tests adequately assess student reading and mathematics comprehension in the form most likely to yield accurate information regarding student achievement of reading and mathematics skills.
>
> (b) REPORT TO CONGRESS—The National Academy of Sciences shall submit a written report to the White House, the National Assessment Governing Board, the Committee on Education and the Workforce of the House of Representatives, the Committee on Labor and Human Resources of the Senate, and the Committees on Appropriations of the House and Senate not later than September 1, 1998 (Heubert et al., 1999, p. xviii).

This was the first time an effort was being made to scientifically prove how testing should be done. It was a momentous occasion as this was a result of a legislative and executive branch request. There had been other studies done by independent organizations but none at the behest of the U.S. national government.

The committee did feel the need to dedicate a chapter to the Legal Frameworks of testing, "In terms of the committee's congressional mandate, the law constitutes one set of norms relevant to whether existing or new tests are used in a discriminatory manner or inappropriately for student promotion, tracking, or graduation" (Heubert et al., 1999, p. 50). This shows that the committee was aware of the challenges that had been made in the courts to high-stakes testing.

With Clinton's proposal of VNT's the committee felt there needed to be analysis how the courts were viewing testing. They found three trends in the testing legal challenges

> The outcomes of some cases depend on whether the decision to administer a high-stakes test is based on a present intent to discriminate. Other cases depend on whether a test carries forward or preserves the effects of prior illegal discrimination. A third claim, grounded in federal civil rights statutes and accompanying regulations, employs an 'effects test' that considers whether a high-stakes test has a disproportionate, adverse impact; whether the use of a test having such an impact can be adequately justified on educational grounds; and whether there are equally feasible alternative tests that have less disproportionate impact (Heubert et al., 1999, p. 52).

The committee analyzed several landmark cases in each of these areas and found the courts focusing on different aspects in trying to rule justly on testing cases. Many of the cases had to do with a special segment of students suing for their constitutional or states' rights.

In the first section of laws that the committee looked at titled, "Claims of Intentional Discrimination"(Heubert et al., 1999, p. 52) the committee summarizes that most of these claims were based on the fourteenth amendment of the U.S. Constitution under the equal protection clause. They committee found,

> …the courts' concern with tracking, remediation, and special education is plainly focused on whether or not students will receive enhanced and effective educational opportunities as a result of the educational intervention. Furthermore, complying with relevant professional testing standards reduces the risk of legal liability for high-stakes assessments (Heubert et al., 1999, p. 55).

It seems the committee found that the courts were acquiescing to the policymakers and the professionals closer to the education of the students. As long as the policies were trying to further education the courts were backing many of the defendants in these cases. It also appears that the committee found the courts wanting to stay out of setting educational policy through court decisions as long as professional standards were being followed.

The second area addressed by the committee had to do with "Claims That Tests Preserve the Effects of Prior Discrimination" (Heubert et al., 1999, p. 55). These cases mostly had to do with school systems coming to grips with the *Brown v. Board* decision of desegregating schools. The courts had a much easier time deciding in these cases since there was Supreme Court backing on non-discrimination in schools. "…claims of this nature are increasingly rare, if only because there are fewer children each year who can show that they themselves attended illegally segregated schools" (Heubert et al., 1999, p. 57).

The final area the committee examined seemed to have more opportunity for future cases. "Claims of Disparate Impact" (Heubert et al., 1999, p. 57) had to do with Title VI of the Civil Rights Act of 1964 which prohibits discrimination on the basis of race, color, or national origin, including limited English proficiency and Title IX of the Education Amendments of 1972 states that it forbids sex discrimination. However, federal regulations call for a higher standard;

> Federal regulations go further: they provide that a federal fund recipient may not 'utilize criteria or methods of administration which have the effect of subjecting individuals to discrimination.' In interpreting this Title VI regulation and similar regulations under Title IX, courts have drawn on interpretations of a federal employment discrimination statute, Title VII (Heubert et al., 1999, p. 58).

The legal terminology for relying on this type of argument is disparate impact claim. The committee went on to explain that after meeting the standards under disparate impact,

> under a disparate impact standard, legal liability may depend in part on whether the test raises problems of measurement, which may be the case if the test has not been validated for the particular purpose for which it is being used or has not been validated for all parts of the test taking population (Heubert et al., 1999, p. 60).

At first glance, this may seem like an opportunity to legally undermine many testing policies and implementations however the legal standards that must be met are cumbersome and are potential dead ends for litigation. In essence, plaintiffs must show that the tests violate some aspects of testing and prove there is a viable alternative in order to be awarded relief from the current testing policy. The fight under disparate impact can be a complex drawn out case which many plaintiffs may be lacking the financial resources to prolong a legal battle. By the time a disparate impact claim is decided, the relief sought could be moot.

After analyzing several aspects of testing the committee does make some general recommendations about testing:

> Recommendation: Accountability for educational outcomes should be a shared responsibility of states, school districts, public officials, educators, parents, and students. High standards cannot be established and maintained merely by imposing them on students.
>
> Recommendation: If parents, educators, public officials, and others who share responsibility for educational outcomes are

to discharge their responsibility effectively, they should have access to information about the nature and interpretation of tests and test scores. Such information should be made available to the public and should be incorporated into teacher. education and into educational programs for principals, administrators, public officials and others.

Recommendation: A test may appropriately be used to lead curricular reform, but it should not also be used to make high-stakes decisions about individual students until test users can show that the test measures what they have been taught.

Recommendation: Test users should avoid simple either-or options when high-stakes tests and other indicators show that students are doing poorly in school, in favor of strategies combining early intervention and effective remediation of learning problems.

Recommendation: High-stakes decisions such as tracking, promotion, and graduation should not automatically be made on the basis of a single test score but should be buttressed by other relevant information about the student's knowledge and skills, such as grades, teacher recommendations, and extenuating circumstances.

Recommendation: In general, large-scale assessments should not be used to make high-stakes decisions about students who are less than 8 years old or enrolled below grade 3.

Recommendation: All students are entitled to sufficient test

> preparation so their performance will not be adversely affected by unfamiliarity with item format or by ignorance of appropriate test-taking strategies. Test users should balance efforts to prepare students for a particular test format against the possibility that excessively narrow preparation will invalidate test outcomes.
>
> Recommendation: High-stakes testing programs should routinely include a well-designed evaluation component. Policymakers should monitor both the intended and unintended consequences of high-stakes assessments on all students and on significant subgroups of students, including minorities, English-language learners and students with disabilities. (Heubert et al., 1999, pp. 278-281).

These were to be the general guidelines for the VNT. The committee also had several specific recommendations that had to do with the categories of "Appropriate Uses of Tests in Tracking, Promotion, and Graduation" (Heubert et al., 1999, p. 282). As one can see these general and specific scientific guidelines were more than enough for any educational test maker to craft clear, coherent and fair tests.

The Use of Tests When Making High-Stakes Decisions for Students: A Resource Guide for Educators and Policy Makers, U.S. Department of Education, Office of Civil Rights, Draft December 8, 1999

This document is more a synopsis of previous research than a researched piece on what testing should be. Then U.S. Department of Education (DOE) Secretary Richard Riley and the

Assistant Secretary of the Office of Civil Rights (OCR) Norma Cantú wrote this guide to educators and policymakers using previously published guidelines and standards.

The guide points out how the DOE and/or OCR support standards that have been printed. "The resource guide is intended to reflect existing test measurement and legal principles" (*The use of tests when making high-stakes decisions for students : a resource guide for educators and policy-makers / U.S. Department of Education, Office for Civil Rights*, 2000, p. i). While this guide does point out research that has been done the guide does dedicate a chapter to legal implications.

Chapter 2 Legal Implications focuses mainly on discriminatory law and cases against special education students or English language learners. The guide does give support to both due process and equal protection practices.

> In the due process section the guide outlines the questions that the courts have addressed; (1) Is the testing program reasonably related to a legitimate educational purpose? (2) Have students received adequate notice of the test and its consequences? (3) Are students actually taught the knowledge and skills measured by the test? (*The use of tests when making high-stakes decisions for students : a resource guide for educators and policy-makers / U.S. Department of Education, Office for Civil Rights*, 2000, pp. 65-67).

The guide illustrates that the seminal cases brought to the courts by plaintiffs such as *Debra P*. need to be followed and supported. It is in this due process section that Riley, Cantú et. al. reflect on how the courts have allowed educational policymakers to govern without the courts getting in the way "Federal courts typically defer to educators' policy judgments regarding the value of legitimate educational benefits sought from the testing programs" (*The use of tests when*

making high-stakes decisions for students : a resource guide for educators and policy-makers / U.S. Department of Education, Office for Civil Rights, 2000, p. 65).

American Educational Research Association

The AERA (American Educational Research Association) by itself issued a position statement on high-stakes testing just before the NCLB was passed. AERA seemed to have recognized that high-stakes testing was becoming more popular with legislators and seemed to want to lay the groundwork for how students should be assessed. In their statement AERA cited their reasons for existence

> The nation's largest professional organization devoted to the scientific study of education. The AERA seeks to promote educational policies and practices that credible scientific research has shown to be beneficial, and to discourage those found effects ("Position Statement of the American Educational Research Association concerning High-Stakes Testing in PreK-12 Education," 2000, p. 24).

This gives the scientific credibility to their research about high-stakes testing.

In their position paper, AERA listed several ways to make sure that assessments were done properly and fairly. Following are the different sections of the paper and the salient synopses of the section:

Protection against High-Stakes Decisions Based on a Single Test

> When there is credible evidence that a test score may not adequately reflect as student's true proficiency, alternative acceptable means should be provided by which to demonstrate attainment of tested standards.

Adequate Resources and Opportunity to Learn

It must be shown that the tested content has been incorporated into the curriculum, materials, and instruction students are provided before high-stakes consequences are imposed for failing examination.

Validation for Each Separate Intended Use

Tests valid for one use may be invalid for another

Full Disclosure of Likely Negative Consequences of High-Stakes Testing Programs

Where credible scientific evidence suggests that a given type of testing program is likely to have negative side effects, test developers and users should make a serious effort to explain these possible effects to policymakers.

Alignment between the Test and the Curriculum

Because high-stakes testing inevitably creates incentives for inappropriate methods of test preparation, multiple test forms should be used or new test forms should be introduced on a regular basis, to avoid a narrowing of the curriculum toward just the content sampled on a particular form.

Validity of Passing Scores and Achievement Levels

Once the purpose is clearly established sound and appropriate procedures must be followed in setting passing scores or proficient levels. Finally, validity evidence must be gathered and reported, consistent with the stated purpose.

Opportunities for Meaningful Remediation for Examinees Who Fail High-Stakes Tests

There should be sufficient time before retaking the test to assure that students have time to remedy any weaknesses discovered.

Appropriate Attention to Language Differences among Examinees

If English language learners are tested in English, their performance should be interpreted in the light of their language proficiency. Special accommodations for English language learners may be necessary to obtain valid scores.

Appropriate Attention to Students with Disabilities

In testing individuals with disabilities, steps should be taken to ensure that the test score inferences accurately reflect the intended construct rather than any disabilities and their associated characteristics extraneous to the intent of the measurement.

Careful Adherence to Explicit Rules for Determining Which Students Are to be tested

There must be explicit policies specifying which students are to be tested and under what circumstances students may be exempted from testing. Reporting of test score results should accurately portray the percentage of students exempted.

Sufficient Reliability for Each Intended Use

It must be shown that scores reported for individuals or for schools are sufficiently accurate to support each intended interpretation.

Ongoing Evaluation of Intended and Unintended Effects of High-Stakes Testing

The governmental body that mandates the test should also provide resources for a continuing program of research and for dissemination of research findings concerning both the positive and the negative effects of the testing program ("Position Statement of the American Educational Research Association concerning High-Stakes Testing in PreK-12 Education," 2000, pp. 24-25).

Joint Committee on Testing Practices

This very prestigious organization (JCTP) has looked at testing since the 1980's. This organization comprises:

> representatives of the National Council on Measurement in Education, as well as the American Counseling Association, American Educational Research Association, American Psychological Association, American Speech-Language Hearing Association, National Association of School Psychologists, and National Association of Test Directors ("Code of Fair Testing Practices in Education (Revised)," 2005, p. 2).

The purpose of the JCTP is "A means by which professional organizations and test publishers can work together to improve the use of tests in assessments and appraisal"("Code of Fair Testing Practices in Education (Revised)," 2005, p. 2).

The JCTP addresses the issue of the general fairness of testing before giving specific guidelines,

> Fairness implies that every test taker has the opportunity to prepare for the test and is informed about the general nature and content of the test, as appropriate to the purpose of the test. Fairness also extends to the accurate reporting of individual and group test results. Fairness is not an isolated concept, but must be considered in all aspects of the testing process ("Code of Fair Testing Practices in Education (Revised)," 2005, p. 2).

JCTP specifically addresses four main areas that should be addressed when undertaking standardized testing " A. Developing and Selecting Appropriate Tests B. Administering and Scoring Tests C. Reporting and Interpreting Test Results D. Informing Test Takers" ("Code of Fair Testing Practices in Education (Revised)," 2005, p. 4). In each section JCTP addresses guidelines for Test Developers and Test Users. These very specific strategies let these groups understand succinctly how tests should be developed and used.

While the JCTP could be a lesson book on how to prepare, implement and use tests the committee cautions,

> The *Code* is not intended to be mandatory, exhaustive, or definitive, and may not be applicable to every situation. Instead, the *Code* is intended to be aspirational, and is not intended to take precedence over the judgment of those who have competence in the subjects addressed ("Code of Fair Testing Practices in Education (Revised)," 2005, p. 4).

Legal and Legislative History

U.S. Constitution

One needs to have the founding articles of law as the basis for any legislation in a democratic society. This is exemplified in the U.S. Constitution which states:

> This Constitution, and the Laws of the United States which shall be made in pursuance thereof; and all treaties made, or which shall be made, under the authority of the United States, shall be the supreme law of the land; and the judges in every state shall be bound thereby, anything in the constitution or laws of any state to the contrary notwithstanding (U.S. Const. Art. VI).

This shows that the founding fathers knew the new United States of America needed a foundational document that could not be subverted. This document was to be the unchallenged legal document that all others would conform to. The U.S. Constitution became the basis for the laws of this newly founded country.

The U.S. Constitution has been held as the benchmark which all laws must be measured against since this document was enacted over two centuries ago. In areas governed by the Constitution, every state or local law must be subordinate to the Constitution. It is this basis that

all laws must eminate from and subscribe to or be judged as constitutional and legal. Without this base document our laws would be baseless and unenforceable.

This above mentioned part of the Constitution specifically states that judges must abide by the provisions made in the Constitution. Were judges allowed to circumvent the Constitution they would be circumventing the foundation of law for this country. The Constitution, as enforced, makes sure this does not happen.

There are provisions within the U.S. Constitution that come into contention when looking at high-stakes testing; "Nor shall any State deprive any person of life, liberty, or property, without due process of law; nor deny to any person within its jurisdiction the equal protection of the laws" (U.S. Const. Amend. XIV, Sect. 2). These few words have been broken down and used to contest the effects of high-stakes testing.

In essence what due process means and how the courts try to interpret this provision is; "Determining what process is due, courts look to the nature of the interest at stake to assess whether the interest is protected and whether the government abused its power in acting to restrict it" (O'Neill, 2003, p. 13). Dayton(2013) stated: "Due process is the framework and lifeblood of any legitimate legal system, establishing substantive standards for fundamental fairness and a procedural system for conducting impartial hearings" (page 92). This reinforces the idea that all governing bodies need to follow the U.S. Constitution or their decisions could be declared unconstitutional.

Before 1970 due process was viewed by the courts as a general deterrent against unfair government processes. However, Quigley (2000) points out "contemporary due process interpretation has been narrowed" (Quigley, 2000, p. 290) and cites the "1972 decision of *Board of Regents of State Colleges v. Roth*. Roth stands for the general proposition that due process is

not a generalized constitutional right to be treated fairly by the government" (Quigley, 2000, p. 290). After this decision by the courts due process was changed, "People who seek to challenge governmental action under the due process clause must first demonstrate to the court they have a constitutionally protected liberty or property interest" (Quigley, 2000, p. 290). This narrowing of due process by the courts allowed high-stakes testing opponents to bring to court cases citing relief from unconstitutional violations of due process.

This following of a process in a timely manner would be classified by O'Neill(2003) as Procedural Due Process, which he explains the courts have viewed as "Individuals are entitled to adequate notice and an opportunity to be heard before governmental deprivation of their constitutionally recognized interest in property or liberty" (page 14).

Under Procedural Due Process O'Neill (2003) categorizes Property Interests and Liberty as protected rights which can be argued in relation to high-stakes testing. For Property Interests he observes the courts have defined these rights as,

> …students have a protected liberty interest in school settings and, in particular, a protected liberty interest in avoiding the sorts of damaging stigma and curtailed career opportunities that can result from the improper implementation of high-stakes exams (page 14).

Substantive Due Process Claims are another area that the courts have used to rule in high-stakes testing cases. "Courts have determined that the Due Process Clause requires that the government avoid taking action that is arbitrary, capricious, does not achieve a legitimate state interest, or is fundamentally unfair" (O'Neill, 2003, p. 19).

There are also equal protection claims that are judged by this phrase within the Constitution.

> The Fourteenth Amendment specifically prohibits discrimination by state governements, requiring states to provide due process and equal protection of the laws to all persons. Equal protection of the laws prohibits differential treatment based on factors that are legally irrelevant, and are instead the products of irrational prejudice or discrimination (Dayton, 2013, p. 131).

Essentially the equal protection clause is mandating that all citizens are treated equally under the law and are not irrationally discriminated against as a class of citizens. "The Equal Protection Clause is regularly invoked to guard against arbitrary classifications that discriminate against a particular group. In order to be in compliance with the clause, all laws that classify citizens must bear at least some rational relationship to a legitimate state interest" (O'Neill, 2003, p. 10).

Brown v. Board of Education of Topeka, 347 U.S. 483 (1954)

The issue in this landmark case was whether excluding minority race children from attending schools attended by the majority race was unconstitutional under the equal protection clause of the Fourteenth Amendment. The opinion of Chief Justice Warren explains the plaintiffs' arguments "The plaintiffs contend that segregated public schools are not 'equal' and cannot be made 'equal' and that hence they are deprived of the equal protection of the laws" (McKay, 1958, p. 204).

The rule of law in this case was the Fourteenth Amendment under which all U.S. citizens had equal protection. Subsequent plaintiffs and advocates argued segregation was a violation of equal protection as the housing, curriculum and resources available to minority students were unequal and therefore unconstitutional. "The action sought a declaratory judgment that the statute and the practices of the defendants were unconstitutional in that they denied the Negro

children the equal protection of the laws guaranteed by the fourteenth amendment" (Wilson, 1964, p. 516).

Applying the rule of law the Court reversed the lower court decision and declared that under the Fourteenth Amendment a segregated school is unconstitutional. "Racial segregation in the public schools of Topeka and elsewhere denied to the segregated groups the equal protection of the laws assured by the fourteenth amendment" (Wilson, 1964, p. 523).

Many have drawn conclusions from this case including the observation that this decision by the Supreme Court changed forever the educational landscape. No longer could students attend separate schools based on their color. While this was a hard pill for many to swallow in the U.S. it changed a perception and a consciousness of a nation.

> With the perspective of hindsight we can always say that one side was right and one was wrong. Where social issues are involved, one side is usually more popular than the other-more in tune with prevailing ideas. But, win or lose, the advocate who presents his position as forthrightly and vigorously as he can performs a service to the court and to society (Wilson, 1964, p. 524).

This shows that not only was this a momentous decision for education advocates but it illustrates how the judicial system must operate in order for issues such as these to be hashed out despite the societal implications.

Students were to be educated without racial segregation said the highest court in the land and it forced the U.S. citizenry to face other areas where there were unconstitutional attitudes and laws that were rooted in irrational treatment based on skin color. *Brown* demonstrated that education can be a catalyst for changing a nation.

Most of the legislation and many court cases that followed in education have direct and indirect ties back to this historic case. The U.S. Supreme Court declared racial segregation de jure unconstitutional yet racial segregation continued de facto throughout the United States.

> For a period over 50 years, educators and scholars have crafted policies to legally desegregate the nation's schools among racial, social class and financial lines. The historic and unanimous case of *Brown v. Board of Education* (1954) made strides against the 'separate but equal' schooling; however, local school district politics blunted implementation efforts over the next 10 years. State and local opposition to the Court's decision gave rise to *Brown II* (1955), which mandated that schools desegregate 'with all deliberate speed'. This, in turn, ultimately enforced mechanisms for desegregated schooling in the United States. (Thompson-Dorsey, Williams, & Houck, 2013, p. 572).

Many cases evolved from the *Brown* decision as many sought to test the limitations of this momentous pronouncement. The cases and legislation that follow are indicators that *Brown* influenced how legislators and judges alike specifically and subliminally considered the greater concepts of this case; equal education for all.

Elementary and Secondary Education Act

The Elementary and Secondary Education Act of 1965 (ESEA) was the next foray for policymakers to try and make education available to all students. As the notion of equal rights for all Americans was being fought in the streets of the U.S. during the early part of the 1960's, education was tagged as a target as well. Even after the historic case of *Brown v. Board of Education* people knew that education was one area that access was not equal and the ESEA sought to address these inequities. President Lyndon Johnson signed the ESEA into law in

1965, "which established the Title I program providing federal funds for disadvantaged schoolchildren" (Vinovskis, 2009, p. 11). As with many ideas that came out of this decade the ultimate goals of ESEA were arguably utopian. "Policymakers in the 1960's hoped the ESEA…would enable the United States by the 1970's to eliminate poverty and provide equal opportunity for all children" (Vinovskis, 2009, p. 11). While the goals of the ESEA were well intentioned, many abused the generosity of these federal funds. Title I funds were distributed without many checks and balances and there were later found out to be vast differences in how Title I funds were spent from district to district. As Vinovskis (2009) stated:

> Rather than supporting a proven set of specific interventions for the improvement of education, the Title I compensatory education program operated more like a general funding mechanism…critics pointed to limited improvements and objected some of the money was being diverted from serving disadvantaged students to other uses (p.11).

The distrust between the different levels of education governance was also arguably aggravated with the ESEA. Local districts were distrusted by the states to spend the federal monies effectively. "As the Johnson administration drafted the ESEA, concerns were raised about whether most state education agencies (SEAs) were capable of distributing and monitoring Title I funds to local school districts" (Vinovskis, 2009, p. 12). The ESEA represented a major step into the state and local governance of school districts. The federal government continued over the next several decades to influence more local and state educational policies by financing programs with federal education agendas attached. ESEA did not go into specifics as to a testing program but ESEA did pave the way to test students to reconcile how federal funds were being spent. The federal government was disbursing large amounts of money and the federal government wanted to know how and where the funds were being spent. Some may have

interpreted this request for fiduciary responsibility as a reason for testing students to prove monies were helping students.

San Antonio Independent School District v. Rodriguez - 411 U.S. 1 (1973)

At issue was the unequal funding of schools due to public schools being funded by the property taxes collected in a particular area. The plaintiffs brought suit that in the Edgewood Independent School district in Texas the property taxes collected were lower due to the lower income of the residents and this made the available resources versus richer districts unequal.

The rule of law the plaintiffs were citing was a reference to the *Brown v. Board of Education* decision and the equal protection doctrine of the Fourteenth Amendment.

> As in *Brown*, the plaintiffs alleged 14th Amendment equal protection violations, proposing two theories of unconstitutionality—first, they argued that education is a fundamental right and is therefore subject to strict scrutiny under equal protection doctrine, and second, they argued that the poor constitute a suspect class, another determination which would trigger strict scrutiny by the Court (Saleh, 2011, p. 103).

Plaintiffs were pleading that due to the unequal amounts of funding available that was a violation of the *Brown* decision that education is a fundamental right. The argument was also being made that the Fourteenth Amendment of the Constitution was being violated because the earmarked education funds were not equal and in the plaintiffs' interpretation under this amendment all students should have equal funding.

The Court applied the fact that there was not a provision for all students to have equal funding no matter what their locale or economic status. "…the Court held in a 5–4 decision that education was not a fundamental right and that the poor did not constitute a suspect class" (Saleh, 2011, p. 106).

The conclusion was that despite the disputed decision, this case was nonetheless a monumental decision, "…established the precedent for property tax-based education funding programs at the state-level" (Saleh, 2011, p. 99) The Court felt that it was up to the legislature of Texas to adopt a better formula for funding schools. In the opinion of the U.S. Supreme Court, this inequity in funding was not due to a purposeful predetermined practice it was one of happenstance. Due to the seeming contradiction to the *Brown* case, it is an issue that continues to be argued,

> The case is made that education finance equality is an imperative of federal civil rights, and the Supreme Court may have a role in ending the ongoing policy decisions by state legislatures, which contribute to a system of public education funding that promotes inequality between wealthier and poorer districts (Saleh, 2011, p. 100).

Scheelhaase v. Woodbury Central Community Sch. Dist., 488 F. 2d 237(8th Cir. 1973)

In 1973 the Northern Iowa District court was to decide whether the firing of Ms. Scheelhaase was lawful, a non-tenured teacher, on the grounds that her students were not being successful on state standardized tests. Ms. Scheelhaase contended as quoted by the court, "invoke[ing] the jurisdiction of this Court pursuant to Amendments I and XIV of the Constitution of the United States"("Scheelhaase v. Woodbury Central Community Sch. Dist," 1973, p. 238). Ms. Scheelhaase was contending her dismissal violated her due process rights assured her by the U.S. Constitution. The district court originally ruled in favor of Ms. Scheelhaase however the verdict was overturned on appeal by the Eighth Circuit Court of Appeals. The rule of law the Appeals Court cited was, "…lower court incorrectly heard the case because the teacher received all the due process to which she was entitled under Iowa state law

and no federal constitutional issue existed". (Crisafulli, 2006, p. 631). The Appeals Court also felt,

> …the Court stated that the District Court should not have interfered with the standards by which the school determined that Ms. Scheelhaase should be dismissed… The Court noted that the termination was valid because the standard upon which the termination was based-the failure of Ms. Scheelhaase students to score at a certain level on a state standardized test-was neither arbitrary nor capricious (Crisafulli, 2006, pp. 631-632).

This was a judgment by an appeals court on a non-tenured teacher but it did set guidelines for later decisions.

Minimum Competency Tests (MCT)

In the 1970's Minimum Competency Tests were the states' first endeavors to have some type of assessment of what were to be the basic skills every student should learn: "This reform came in response to complaints, from many segments of a society, that there need to be higher standards in schools" (Christie & Casey, 1983, p. 31). American society was becoming increasingly alarmed that students did not have the basic skills necessary to perform in the work place. "In an effort to blunt fears that social promotion policies, unfocused curricula, and diluted academic standards combine to devalue the high school diploma, States began to implement MCTs"(Heise, 2009, p. 2).

There was also concern that the ESEA while providing resources did not measure whether the resources were being utilized effectively or to the benefit of students "In response to disillusionment with educational reform policies focused on 'inputs' (better resources, better curricula, new teaching methods), policymakers shifted attention to interventions that focused on outcomes" (Haertel & Lorie, 2004, p. 15).

This was the first step into standardized, high-stakes tests as many of these tests were tied to promotion or graduation, "some state legislatures have passed laws that require school districts to adopt minimum competency standards and to deny graduation to students who do not meet these standards" (Christie & Casey, 1983, p. 31). This was troublesome as far as the law was concerned in that the implementation of standards, the testing of these standards and the denial of diplomas was happening simultaneously. "Most states found it far easier to enact MCT legislation than to implement the tests. Resistance to MCTs quickly emerged due to the legal and political fallout incident to students' failing MCTs and, in particular, not graduating" (Heise, 2009, p. 2).

> While the denial of the diploma has a certain deterrent value, its application in the instant case would be analogous to asserting that the immediate and indefinite incarceration without trial of an individual upon suspicion of the commission of a crime would have a deterrent effect on other potential offenders (Quigley, 2001, p. 17).

The consequences were; fail the test and be denied a diploma along with the social stigma of being viewed as a failure, despite accomplishments in the classroom. Many took offense to this new type of education reform. "No matter how helpful high-stakes testing is in prodding, pushing and pulling needed educational reform, it cannot be allowed to proceed at the price of giving up the constitutional due process rights of children" (Quigley, 2001, p. 17). Since minor children are viewed as having little to no rights it is up to other educational stakeholders to speak up in their behalf. It has been left up to the lawmakers to be the legislative experts on what is wrong with education and how to fix it.

As the results from MCT's were made public many had issues with holding students to even higher standards. The idea was that students should not just be achieving minimums in order to graduate, they needed to excel.

> …there is little evidence that MCT's were the reason for improvements on other examinations; and the improvements in passing rates on MCT's themselves may have reflected more the effects of drill and practice narrowly focused on testing skills over time, popular concern shifted from an emphasis on 'basic skills' towards more complex 'higher-order thinking' skills (Haertel & Herman, 2005, p. 14).

The high-stakes component of MCT's was that students were to pass a test in order to graduate from high school.

> The approach that caught on was minimum competency testing, an outgrowth of the 'back to basics' movement of the 1970s. The minimum competency test (MCT) was a basic-skills test, usually in reading and mathematics. Typically, students were required to pass the MCT in order to receive a regular high school diploma, although MCTs could be used in other ways (Haertel & Herman, 2005, p. 13).

At stake was that if the student did not pass the test they would not be eligible to receive a diploma until they passed the test. If a student did not pass the test by graduation the student would receive a certificate of attendance or some other type of differentiated high school diploma. These less than standard diplomas caused some educational stakeholders to question the reasoning.

> When students have performed appropriately, when they have done as much as they are capable, they deserve reward. They deserve to graduate. They deserve a high school diploma. They do not deserve a certificate of attendance or a differentiated diploma that

> implicitly or explicitly indicates to all an inferior or subordinate level of performance (Palardy, 1984, p. 404).

Other scholars of the time philosophized,

> A student whose diploma had been withheld on the basis of reading and math scores might well sue the local or state education agency on the grounds that reading and math scores have weak validity in predicting later life success—the criterion on which minimum competency programs are predicated. As things presently stand there is little scientific basis for the idea that we know what the important competencies-are; and the tests presume to test these competencies cannot with-stand serious scrutiny of their ability to predict "competency" in life skills, survival, functioning, or other such things (Haney & Madaus, 1978, p. 8).

Despite these criticisms “By 1980, statewide minimum competency testing requirements had been implemented in 29 states, most having been initiated in 1975 or later” (Haertel & Herman, 2005, p. 14). One could assume that this wave of accountability was placing the stakes squarely on the student. Popham (2001) reflecting on the MCT’s, “The policymakers who installed these competency tests were actually displaying their doubts about public school educators” (page 5). He goes on to support this position by stating;

> As legislators and other educational policymakers saw the problem, if kids were making their way through an educational system without having learned how to read, write, or fill out a job application, then someone was falling down on the job—falling down on the job of teaching. As blame was being assigned, educational stakeholders accepted the emphases of MCT’s. Members of the business community lined up solidly behind the establishment of minimum competency testing programs. If competency programs could

> even partially guarantee that graduates possessed those skills, then corporate America quite naturally endorsed those tests with gusto (Popham, 2001, p. 5).

The perceptions of MCT's were to make sure students were graduating with basic skills necessary to be a productive U.S. citizen. Teachers were to reinforce these ideals by educating students to be able to meet these goals. By the definition of high-stakes tests, MCT's were not high-stakes tests. However, it is a first step towards high-stakes testing since MCT's raise the expectations for teachers and students.

Policymakers did not make any distinction between students who would have to take the MCT, therefore the test makers followed suit. The test makers made tests that in retrospect were socio-economically biased. Some foresaw this bias as an example of one of the inaccuracies a high-stakes test could present. "The principle of disproportionate outcomes is but one of several measures of equal educational opportunity that might be applied to competency testing" (Serow & Davies, 1982, p. 529).

Others saw problems in the how measures implemented would be acted upon. There seemed to be trepidation in trying to measures what should be taught in schools.

> The nature and magnitude of the effects that accountability systems produce in schools depend on the characteristics of the particular system, which, in turn, are conditioned by the political-economic origins and the educational context of the state in which the system is implemented (Carnoy, Elmore, & Siskin, 2003, p. 14).

Therefore, the goal of the standardized high school diploma was seemingly a goal that was never met despite the repeated enforcement of MCT during the 1970's into the mid 1990's. Most of the warnings against MCT's went unheeded. Even the data collected from MCT's was marginally useful at best.

> States initial accountability efforts generally entailed judging schools on the basis of the amounts and kinds of inputs they had—library books, lab equipment, condition of buildings, percent certified teachers, and class size for example (Carnoy et al., 2003, p. 15).

During this time when schools were measured via resources MCT's were enacted without a regard for access to opportunity. "…the causal link between student outcomes and such inputs was not clear once neighborhood income was accounted for" (Carnoy et al., 2003, p. 15). The realism is that "Equal opportunity may require unequal resource allocation, since more resources may be needed to bring one group up to the desired level of outcomes than another" (Serow & Davies, 1982, p. 532). What was happening was the MCT was actually backfiring on the policy makers, which at this time was state level on down to local, who were trying to make sure every graduating student had basic competencies. The results were that students that were low performers continued to perform poorly on poorly designed tests. The goal of making all high school diplomas equal was falling short of the mark. Instead of bringing students up to basic levels of understanding, MCT's seemed to building an even bigger gap between the haves and have not's. Students that were not passing the tests after several attempts were more apt to forego graduation and drop out. Students that were passing the tests were minimally affected by the tests since they possessed more than the basic skills that the tests were assessing. This would not discourage policy makers as we shall see policy makers felt that by ratcheting up the stakes that more students would perform.

Educational observers were pointing out difficulties in leveling the educational comparison field. They surmised that these unequal opportunities were based on unequal resources. Since the *Brown* decision, unequal resource allocation was not acceptable. *Brown*

supported equal resources for all students and MCT's were bringing to light this disparity. The affected groups of students were looking to the courts for relief.

Debra P. v. Turlington, 474 F.Supp. 244 (M.D.Fla.1979)

As the federal and state governments began to impose high-stakes testing there were challenges to the legality of these tests. One of the first landmark cases was *Debra P. vs. Turlington.* The issue was had to do with the legality of administering a test to a student that did not have enough time to prepare for the state (Florida) exit exam. Specifically, "the state legislature of Florida enacted a statute under which high school students had to pass a functional literacy examination to graduate" (Gerber, 2002, p. 3). Debra P. had failed the test and according to the new state laws, she was not allowed to graduate with a high school diploma. The plaintiffs brought suit claiming, according to the rule of law, her rights of due process rights had been violated per the Fourteenth Amendment "Nor shall any State deprive any person of life, liberty, or property, without due process of law" (U.S. Const. Amend. XIV, Sect. 1).

The application of the law that was argued was that the denial of the high school diploma deprived Debra P. of property of her high school diploma without due process. "…the court concluded that the students had a property right to their diplomas and thus, the state should have provided adequate notice before invoking a diploma sanction" (Gerber, 2002, p. 3). This decision provided a framework for other states to follow or face a similar sanction against their high-stakes test. In reflection on the influence of the *Debra P.* case; the argument against the state of Florida provided a cautionary tale to other states "When a state adopts an exit exam policy, it assumes an affirmative duty to provide each student with an opportunity to learn the tested content" (Welner, 2010, p. 98).

Many subsequent challenges to high-stakes testing were a result of some perceived injustice against; minority groups, special education or a combination of both. As stated previously, the focus of this research is the overarching question as to the fairness of high-stakes testing and the legal challenges to it. One could surmise that minority groups and special education would not be at the forefront of these legal cases if somewhere along the line a fairness issue had been brought to light about all students and high-stakes testing.

A Nation at Risk

The true turning point in how classroom assessment changed was when the report A Nation at Risk (ANAR) was released in 1983. This scathing review of the education system in the U.S. focused on how the education system was failing and was dooming a nation.

> Like the many 1980s' reports, A Nation at Risk was a forceful reminder that the nation's longstanding interest in academic achievement and education standards earlier expressed in the Post-World-War-II Era was being renewed. It signaled the ever-growing federal role in public education characterized by an interest in providing and achieving equality of educational opportunity as well as developing citizens capable of performing effectively in the Global Economy (Johanningmeier, 2010, p. 348).

This report touched chastised the American education system. It admonished the past mistakes, "Our society and its educational institutions seem to have lost sight of the basic purposes of schooling, and of the high expectations and disciplined effort needed to attain them"(National Commission on Excellence in, 1983, p. 2). ANAR used Cold War insecurities to fan the flames of outrage at the state of education,

> If an unfriendly foreign power had attempted to impose on America the mediocre educational performance that exists today, we might well have viewed it as an act of war.

> As it stands, we have allowed this to happen to ourselves. We have, in effect, been committing an act of unthinking, unilateral educational disarmament (National Commission on Excellence in, 1983, p. 2).

These inflammatory statements got the attention of educational stakeholders who may have ignored this as another government report.

ANAR also makes several veiled attacks (and one very pointed attack) against the Minimum Competency Tests that had been used up to this point to the consternation of the Committee,

> We should expect schools to have genuinely high standards rather than minimum ones. When you give only the minimum to learning, you receive only the minimum in return. In contrast to the ideal of the Learning Society, however, we find that for too many people education means doing the minimum work necessary for the moment, then coasting through life on what may have been learned in its first quarter. But this should not surprise us because we tend to express our educational standards and expectations largely in terms of 'minimum requirements.' 'Minimum competency' examinations (now required in 37 States) fall short of what is needed, as the 'minimum' tends to become the 'maximum,' thus lowering educational standards for all (National Commission on Excellence in, 1983, pp. 8,11,19).

The report cited risks that were being observed and most of these risks had some type of relationship with testing;

- International comparisons of student achievement, completed a decade ago, reveal that on 19 academic tests American students were never first or second and, in comparison with other industrialized nations, were last seven times.

- Average achievement of high school students on most standardized tests is now lower than 26 years ago when Sputnik was launched.
- Over half the population of gifted students does not match their tested ability with comparable achievement in school.
- The College Board's Scholastic Aptitude Tests (SAT) demonstrate a virtually un-broken decline from 1963 to 1980. Average verbal scores fell over 50 points and average mathematics scores dropped nearly 40 points.
- College Board achievement tests also re-veal consistent declines in recent years in such subjects as physics and English.
- Both the number and proportion of students demonstrating superior achievement on the SATs (i.e., those with scores of 650 or higher) have also dramatically declined.
- There was a steady decline in science achievement scores of U.S. 17-year-olds as measured by national assessments of science in 1969, 1973, and 1977.
- Average tested achievement of students graduating from college is also lower (National Commission on Excellence in, 1983, p. 5).

Based on these assessed risks the NCEE made recommendations in several key areas. These recommendations were perceived by the NCEE as immediate fix and the implementation was almost a patriotic necessity;

> In light of the urgent need for improvement, both immediate and long term, this Commission has agreed on a set of recommendations that the American people can begin to act on now… Our recommendations are based on the beliefs that everyone can learn, that everyone is born with an urge to learn which can be nurtured, that a solid high school

> education is within the reach of virtually all, and that life-long learning will equip people with the skills required for new careers and for citizenship (National Commission on Excellence in, 1983, p. 13).

The Committee felt the need to group their recommendations into five key areas. These areas were; Content, Standards and Expectations, Time, Teaching and Leader and Fiscal Support. Under the Standards and Expectations recommendation the Committee alluded to testing in general by stating, "We recommend that schools, colleges, and universities adopt more rigorous and measurable standards. This will help students do their best educationally with challenging materials in an environment that supports learning and authentic accomplishment" (National Commission on Excellence in, 1983, p. 15). ANAR then cites a specific recommendation that will help achieve this goal,

> Standardized tests of achievement (not to be confused with aptitude tests) should be administered at major transition points from one level of schooling to another and particularly from high school to college or work. The purposes of these tests would be to: (a) certify the student's credentials; (b) identify the need for remedial intervention; and (c) identify the opportunity for advanced or accelerated work. The tests should be administered as part of a nationwide (but not Federal) system of State and local standardized tests. This system should include other diagnostic procedures that assist teachers and students to evaluate student progress (National Commission on Excellence in, 1983, p. 15).

A Nation at Risk was not a law but many a state government heeded the call to action and took the recommendations to heart by drafting education policies based on ANAR. Despite the misgivings from the NCEE, there was an increase in states adopting minimum competency tests.

Local school boards were deciding the most basic concepts that students needed to understand. What was clear was that ANAR had refocused the all educational stakeholders' efforts towards a quality education. The problem was, getting a consensus on how this was to be done and if those efforts were fair and legal.

America 2000: An Education Strategy

America 2000 was introduced by President George H. Bush in 1991 as a result of a successful meeting at the Charlottesville Education Summit. The Summit consisted of an invitation only gathering by President Bush of White House staff and governors from thirteen states. This put the White House at odds with Congress who felt slighted by not having input on federal education policy. The Summit leaders crafted America 2000 with clear ideas on what America 2000 was, "AMERICA 2000 is a national strategy, not a federal program"(W. D. C. Department of Education, 1991, p. 11). Despite being at odds with Congress over this meeting America 2000 tried to focus the nation on what should be done in education.

> President George H. Bush's announcement of the six national education goals and the creation of the National Education Goals Panel (NEGP) refocused attention on the problems of educational reform in America. However the question remained as to what general strategies and specific actions should be pursued in order to achieve those objectives and whether the presence of national education goals would influence how educators and political leaders crafted reform policies (Vinovskis, 2009, p. 32).

First, there were four main strategies that President Bush and his colleagues outlined needed to be understood as the general outline for the U.S. to succeed as an education nation.

> Our vision is of four big trains, moving simultaneously down four parallel tracks: Better and more accountable schools; a New Generation of American Schools; a Nation of

> Students continuing to learn throughout our lives; and communities where learning can happen (W. D. C. Department of Education, 1991, p. 13).

Under these strategies were six goals that were introduced as goals for what the states should be working towards to further education in America.

> 1. All children in America will start school ready to learn.
>
> 2. The high school graduation rate will increase to at least 90%.
>
> 3. American students will leave grades four, eight and twelve having demonstrated competency in challenging subject matter including English, mathematics, science, history, and geography; and every school in America will ensure that all students learn to use their minds well, so that they may be prepared for responsible citizenship, further learning, and productive employment in our modern economy.
>
> 4. U.S. students will be first in the world in science and mathematics achievement.
>
> 5. Every adult in America will be literate and will possess the knowledge and skills necessary to complete in a global economy and exercise the rights and responsibilities of citizenship.
>
> 6. Every school in America will be free of drugs and violence and will offer a disciplined environment conducive to learning (W. D. C. Department of Education, 1991, p. 19).

It was within the introduction of America 2000 that President Bush specifically addressed the strategy of standardized testing, "We will develop voluntary—let me repeat it—voluntary national tests for 4^{th}, 8^{th} and 12^{th} graders in the five core subjects" (Vinovskis, 2009, p. 42). America 2000 specifically addresses testing,

> These American Achievement Tests will challenge all students to strive to meet the world class standards and ensure that, when they leave schools, students are prepared for further

study and the work force. The tests will measure higher order skills (i.e. they will not be strictly multiple choice tests) (W. D. C. Department of Education, 1991, p. 48).

The bigger issue within America 2000 was a provision for setting up new schools and school choice. Many legislators did not warm to the idea of public monies being steered towards private schools. Bush wanted to set up 535 new schools structured in America 2000 goals and aspirations. Congress already slighted by not being invited to the Summit vetoed any federal monies being spent on new public schools or private schools.

Therefore, America 2000 did not pass Congress despite the efforts of President Bush to reaffirm his position as the 'Education President'. "Overall, the 102nd Congress did little legislatively to advance the Bush administration's America 2000 program" (Vinovskis, 2009, p. 49).

Within the America 2000 programs were some ideas that were developed. There was a feeling that there needed to be some type of accountability on a national scale. Specifically, there was the desire for researching a national test for the common core subjects. From the National Education Goals Panel of 1991(NEGP), Congressional leaders negotiated to investigate some type of national tests by forming the National Council on Educational Standards and Testing (NCEST), a most unfortunate acronym for a very important group. The recommendations by this esteemed group of bipartisan participants was "There should be a national assessment system rather than a single national examination for each grade and subject tested…the council called for multiple assessments rather than a single test"(Vinovskis, 2009, pp. 53,54). While many policymakers could not find fault with a bipartisan report many of those who would be affected directly were not as welcoming. "About 50 prominent educators and researchers called for 'hitting the brakes' on efforts to create a national test. These individuals protested, 'We believe

that the pursuit of such standards does not require—and could be severely compromised by—a [single] test'" (Vinovskis, 2009, p. 54).

The issue of testing under America 2000 was more than just what type of test to give. America 2000 brought to the forefront the idea of the separation of public funds from private schools. Congress voiced their displeasure with America 2000 by turning down funding for private schools. At the local level, the issue of funding for private schools was an ancillary issue. Some were starting to question the fairness of private schools being exempt from states education testing requirements for public school students for promotion or graduation.

Rankins v. Louisiana State Bd. of Elementary & Secondary Educ., 637 So. 2d 548, 551 (La. App. 1st Cir. 1994)

This issue in this case had to do with the constitutionality of allowing private schooled students in the State of Louisiana to forgo the exit exam Louisiana Graduate Exit Exam (GEE) instituted in the public schools by the BESE (Board of Elementary and Secondary Education). Students that brought this case to court had failed the state test and thus denied a high school diploma. These public school students felt this was unjust since private school students in the same state did not have to pass the state test in order to receive a diploma.

> The plaintiffs maintained that although they each failed all or part of the GEE, they successfully completed the state's requisite standards for graduation (i.e., twenty-three Carnegie units) and, therefore, were entitled to their high school diplomas."6 These public school students argued that BESE unfairly administered the GEE since their counterparts at state approved non-public schools were not required to take it (Johnson, 1995, p. 188).

The plaintiffs argued the test violated the Equal Protection provision of the U.S. Constitution since not all students being schooled in Louisiana were subject to the state exit exam. Rankins argued students were being treated differently based upon whether they attended public school or private school.

According to the plaintiffs, the state had violated the rule of law, "BESE, The Board of Elementary and Secondary Education, (as an agent of the state) violated their equal protection rights guaranteed by the Fourteenth Amendment of the United States Constitution" (Johnson, 1995). Their complaint was based on the exceptions to passing the GEE which stated,

> Students were not required to take the GEE prior to receiving their diplomas: (i) state approved non-public school students; (ii) students who participate in home study programs; (iii) students who receive the General Educational Development Diploma ("GED"); and (iv) state approved non-public school students who subsequently re-enter the public school system at the eleventh or twelfth grade levels (Johnson, 1995, p. 188).

While the district court agreed with the plaintiffs, the defendants successfully appealed the case and were victorious in setting aside the original decision. "…the Louisiana First Circuit Court of Appeal reversed, finding that the GEE was constitutionally administered"(Johnson, 1995, p. 189). This application of law by the Court was based on leaving the administration of testing up to the individual states. The Court was of the opinion that how private schools conduct themselves versus public schools are allowed to be different. It is apparent that the Court felt there was no unconstitutionality in Louisiana with the exceptions to taking the GEE.

The conclusion in this case, is that while there are challenges to the legislation of high-stakes testing clearly there can be exceptions to applications of the test. Louisiana showed that public schools are very much under the jurisdiction of state mandated tests while private schools

are not. This case showed the separation between applying public laws to private institutions in the case of education does have a legal basis.

Goals 2000: Educate America Act

The next legal easing into the education field by the federal government was by the Goals 2000 legislation heralded by then candidate Bill Clinton who in 1994 wanted education to adhere to goals. These purpose of this legislation were very similar to the goals Clinton helped craft at the Charlottesville Summit. Goals 2000 introduced its' purpose as,

> An Act to improve learning and teaching by providing a national framework for education reform; to promote the research, consensus building, and systemic changes needed to ensure equitable educational opportunities and high levels of educational achievement for all American students; to provide a framework for reauthorization of all Federal education programs; to promote the development and adoption of a voluntary national system of skill standards and certifications; and for other purposes (Congress, 1994, p. 1).

It was made very clear in the legislation that there was not to be standardized testing to evaluate the understanding of concepts by students. In fact in the redrafted Goals 2000 the legislation, "which now included provisions for drafting national service delivery goals and prohibited tests for high-stakes purposes"(Vinovskis, 2009, p. 70). A more explicit provision to avoid high-stakes testing could not have been drafted.

Goals 2000 addresses assessments as an evolving effort to measure what students are learning but questions whether high-stakes should be associated with these tests;

> Successfully monitoring and assessing progress in improving the academic achievement of all students is crucial. Measuring the performance of all students against standards is

> neither simple nor inexpensive. Tests and other assessment measures, like curriculum and instruction, are tools for teaching and learning that serve as gateways to academic achievement. Should high stakes be attached to their results, and if so, for whom? Who should be held accountable for assessment outcomes--teachers, students, schools, or districts--and for what: overall scores, value added progress, other indicators of school level progress? What is the future impact of failing to promote children as a result of poor assessment outcomes? Will sanctions against poor performing schools limit their ability to improve? (*Goals 2000 educate America : the national education goals*, 1993, pp. 27,28).

This shows that as Clinton was asking the states to follow his guidelines there were still numerous questions that needed to be answered about high-stakes testing. Allowing for states to attempt to answer these questions fairly without harming any group of students was an impossible task. The states had to make sure their rulings were fair and legal to all student groups and sub-groups. Assuming that every state and local district would follow federal guidelines without violating any legal standard would be irrational as many of the guidelines were open to interpretation by state or local officials.

The controversy was with the vague descriptions of opportunities to learn (OTL) and the intrusion of the federal government into state mandated territory. OTL standards were to be

> Voluntary standards to address such as: the capability of teachers to provide quality instruction to their areas; the extent to which teachers and administrators have continuing and ready access to the best knowledge about teaching and learning and how to make needed school changes, and the quality and availability of challenging curricula geared to meet world class standards (Kosar, 2005, p. 115).

States were not supportive of the opportunities to learn provision, in fact they anticipated pending legal action should they adopt Goals 2000, "Governors, in contrast, were adamantly against being coerced by the federal government to adopt OTL standards…Some of the more obvious problems were: Any official measure of an OTL would set up states to be sued for the failure to meet these standards…" (Kosar, 2005, p. 119).

Eventually with bipartisan negotiation and support Goals 2000 did eventually pass but it lacked the teeth of most mandates in that it was considered to be voluntary for the states to follow.

Once the 1994 elections were over and the congress changed to a Republican majority Goals 2000 was earmarked as a Democratic idea that needed to be repealed or at worst unfunded. The year 2000 was when the educational goals were to be reached and Goals 2000 to be considered for reauthorization however, "When the year 2000 came, Goals 2000 expired" (Kosar, 2005, p. 174). The growing partisanship in Congress and the distrust between educational legislators and stakeholders seemed to help with the demise of Goals 2000, "…the growing congressional hostility towards federal attempts to guide state and local education reforms left Goals 2000 with little ability to shape school policies and practice"(Vinovskis, 2009, p. 115). While Goals 2000 was dying a slow death President Clinton was still trying to get other education legislation passed.

IASA

While Goals 2000 was had progressed from the Charlottesville Summit and America 2000, the Improving America's School Act (IASA) was an attempt by President Clinton "to reform the federal government's largest elementary and secondary aid program, Chapter 1 aimed at using federal power to raise education standards, but in a manner less direct than creating

national standards outright" (Kosar, 2005, p. 153). This attempt in September 1993 of the Clinton administration was to reform the Elementary and Secondary Education Act.

The ESEA was landmark legislation when it was introduced by President Johnson in the 1960's. However, much like how many reflect that the Constitution is a living document, many felt the ESEA should adapt to different struggles in education decades after it was adopted. One of the many reasons was that Chapter 1 was an aid package that education leaders used at their discretion without much oversight. This led to mismanagement of funds and little progress attributed to this legislation. Clinton attempted to ride on the heels of his victory with Goals 2000 and looked to the IASA as a way to extend the many of the ideas in Goals 2000. "in the 1994 reauthorization that required states to establish content and performance standards in reading and mathematics and to design assessments aligned with those standards" (McDonnell, 2005, p. 29). The assessments piece of this legislation would prove to be a tripping point for Clinton and his ideas of implementing voluntary national testing.

IASA was crafted by the undersecretary of education at the time Marshall S. Smith. Smith along with Jennifer O'Day had called for systemic reform in a widely praised article written in the book *Designing Coherent Education Policy: Improving the System* edited by Susan Fuhrman. The chapter from Smith and O'Day stated,

> …systemic curriculum reform offers an alternative that might have strong positive effect on equal opportunity. Curriculum frameworks that contain challenging content and that are developed in the spirit of open discussion among diverse groups could provide direction for major curriculum reform for all schools and their students (Fuhrman & Consortium for Policy Research in Education, 1993, pp. 298-299).

The research article of Smith and O'Day considered the many facets of systemic curricular reform, including the legal considerations,

> Courts and legislatures could redefine their understanding of equal educational opportunity in the context of a systemic curricular strategy. Equal opportunity for poor and minority students would be achieved when their schools meet the standards for providing all of their students the opportunity to achieve a high standard of performance on assessments based on the curriculum frameworks. Of course, providing this opportunity may require very different strategies for different students (Fuhrman & Consortium for Policy Research in Education, 1993, p. 290).

Smith helped Clinton craft the IASA to meet goals established by Smith. Congress concurred with these proposed goals and adopted them as part of the IASA:

> Undersecretary Smith thought that Title I should be in the mainstream of states' standards-based reforms in two fundamental way. First, the required standards and assessments in Title I should be uniform for all students. Academic expectations were supposed to be the same for different populations of student, which had not been the case when the program mainly supported remedial instruction (DeBray, 2006, p. 29).

This was historic since this was the first time that all students were to be subjected to the high-stakes test. While this was supposed to be a step in the right direction this became a point of contention in later legislation as special needs students were made to meet the same standards as those without mental or physical challenges.

The IASA proposal of the measuring of students meeting standards by assessments was where the phrase Adequate Yearly Progress (AYP) was coined. This idea that students were progressing from year to year by measurement on a criterion-referenced test was an idea that was

adopted for IASA as well as later federal legislation. There were also to be consequences for schools not meeting AYP under IASA.

AYP was far from being the only new idea enacted under this legislation, some of the other new ideas under IASA were;

- Transform Chapter 1/Title 1 from a pull-out compensatory aid program to a standards-based program.
- Make the standards-based reform the basis for systemic reform.
- Use Title 1 as general aid; that is use it to improve the education of all children by conditioning the acceptance of the aid on states enacting standards-based reform for all schools (Kosar, 2005, p. 159).

These provisions changed the direction of the original Chapter 1. These last three provisions transformed Chapter 1/Title 1 to an aid program available for all students not just poverty stricken or disadvantaged students. Also, learning from past legislation, Clinton and Smith did not push for national standards but did push for the idea that "control over curricula would remain with the states and localities" (Kosar, 2005, p. 159). This allowed Clinton to try and further champion his win for the ideas contained in Goals 2000 as a way to get approval from Congress for IASA.

There were years of contentious debates in both houses of Congress, committees, hearings and subcommittees but with the support of many of key educational supporters such as Sen. Kennedy, IASA was able to pass. Proponents of the law extolled the virtues, "increased flexibility in the uses of funds so that teachers can find the best ways to help children to achieve more" (Jennings, 1998, p. 149). Many remarked how education policy had been battered but IASA emerged as a winner because it was an aid bill that constituents could relate to. IASA

made clear that ESEA funds would be directed towards schools and affect education on the local levels. It also mattered that Senator Kennedy was able to make compromises with key Republican leaders by "adding provisions that mattered to them" (Jennings, 1998, p. 149). This allowed the legislation to pass and both parties could declare political victory. Clinton and the 103rd Congress had been able to capture the titles that President Bush had so desperately wanted, 'Education President and Education Congress'. It was an education policy war that set the stage for the federal government to have a more direct hand in the governance of education.

IASA was now a law but there were provisions that fell to the wayside, one such provision was voluntary national tests. Clinton and his supporters had very little awareness of how to develop much less implement a voluntary national test. While Clinton tried to assuage the American public by making the announcement during his second State of the Union address in 1997:

> Clinton's timing was strategic; if he had proposed the tests prior to the November election, congressional Republicans might well have criticized his wanting to exert excessive control over local school boards and used this as a rationale for cutting programs (DeBray, 2006, p. 34).

The Republicans were set against any type of national test or output and wanted more investment in schools and learning, specifically private school choice.

The National Academy of Science (NAS) was charged with finding out whether a national test, voluntary or not, would address schools shortcomings in education. NAS determined national tests "had an unacceptably high risk of producing scores that inaccurately portrayed student learning" (Kosar, 2005, p. 173). Since accurate assessments were essential to standards reform this national report was the death knell for any type of national testing under

IASA. State tests were still being implemented and challenged in the courts, despite the federal attempts at nationalizing tests.

Erik V., by and through Catherine V. v. Causby, 977 F. Supp. 384 - Dist. Court, North Carolina 1997

The issue in this case was parents objections to the state policy of not allowing promotion based on failing the state tests being implemented. The Johnson County, NC Board of Education passed a policy in June of 1996 outlining promotion criteria "to address a perceived performance deficit on the part of the students of Johnston County Schools" ("Erik V. , by and through Catherine V. v. Causby," 1997). The Board of Education is allowed to pass policy pertaining to school matters. The parents felt this policy violated state and federal laws that protected students from being subjected to this type of promotion criteria. "Plaintiffs contend that the use of end-of-grade tests to make promotion decisions violates various federal and state constitutional and statutory right" ("Erik V. , by and through Catherine V. v. Causby," 1997). The specifics of the Johnston County Board of Education Policy 842 at issue were,

> The policy provides that students in grades three through eight who fail to reach Achievement Level III during the end of year administration of the end-of-grade tests developed by the State of North Carolina will be retained. Students who do not score at Level III on the first administration of the test are provided with a brief remediation and retested. If a student scores only at Level I on the retesting, he is required to attend summer school, which is also optional for those students attaining Level II on the retest. All of those students are retested a third time and are promoted if they score at Level III.

> If a student scores below Level III but has earned A, B, or C grades on grade-level work during the school year, the teacher and principal are required to review the student's work to determine whether he or she is performing at grade level notwithstanding the end-of-grade test scores. If they believe the student is performing at grade level, they must seek a waiver of the policy from a committee of other educators convened for this purpose. If the teacher and principal decide that a waiver would not be appropriate, the parent may appeal to the principal for reconsideration. Teachers were required to notify parents by the end of the first semester if students were at risk, and to offer remediation ("Erik V. , by and through Catherine V. v. Causby," 1997).

According to the plaintiffs, the rule of law was the equal protection clause of the U.S. Constitution along with state constitutional laws of North Carolina. The application of law was that due to the constitutional violation students would be irreparably harmed due to not being promoted to the next grade level. The plaintiffs claimed if retained the students would be delayed year(s) by this retention and suffering fiduciary damages.

> Plaintiffs also attempt to argue that they will suffer irreparable harm if they are retained in grade because they will complete school a year later in the future and thus lose a year of opportunity to begin a career, to attend college, or to start a family. Additionally, Plaintiffs contend that retention is likely to affect them in a variety of other, unquantifiable ways, such as low self-esteem, a negative attitude about school, and a smaller chance of succeeding in school ("Erik V. , by and through Catherine V. v. Causby," 1997).

The conclusion was that the court disagreed with the plaintiffs and sided with the defendants because in the opinion the court ruling for the plaintiffs would be overstepping its boundaries by ruling on academic policy.

> The harms that the Johnston County schools would suffer if this motion were granted are severe and legion. The most severe consequence would be the effect of having a carefully engineered promotion/retention policy superseded by a federal court, outweighing the policy decisions made by an elected school board in a public deliberative process. The administrators and teachers in the Johnston County schools would also suffer a lapse of credibility from the sudden overturning of a policy they have been enforcing in their classrooms, as well as from the disruptive effect of having students deemed unready for promotion being mixed in with others ready for the challenges of a new grade ("Erik V. , by and through Catherine V. v. Causby," 1997).

The court was saying that interceding on behalf of the plaintiffs would be problematic since it would seem the courts would set policy and not the elected officials entrusted to do so.

GI Forum Image De Tejas v. Tex. Educ. Agency, 87 F. Supp. 2d 667, 673 (W.D. Tex. 2000)

In this case of an exit exam given by the state of Texas, at issue was the alleged disparity in passing rates between groups of students. Most the students failing the Texas Assessment of Knowledge and Skills (TAKS) were minority students such as Mexican-American and African-American. Multiple opportunities were given to all students to help them pass the TAKS, however these sub-categories of students were still disproportionally failing this exit exam and thus denied high school diplomas.

The rule of law was laid out by the Mexican American Legal Defense (MALDEF) fund who argued “that Texas's exit exam violated students' equal protection, due process, and

statutory rights" (Heise, 2009, p. 5). MALDEF was clearly stating that their clients were harmed by the TAKS since it seemed they were being unfairly targeted to be denied a high school diploma they had otherwise earned. MALDEF felt this was in violation of their Fourteenth Amendment rights of being able to have the same opportunity as any other group of students to pursue a high school diploma. The due process being disputed was a claim based on the plaintiffs feeling there was not a process to voice their claims and affect relief from their perceived harm.

While this was a common claim made against states using exit exams as final criteria for earning high school diplomas, only one part of these allegations were held up by the court. The Title VI component of statutory rights was the only claim that the court would accept. The opinion of the court was there was enough evidence already provided to further a case (prima facie) against the state for discrimination.

The conclusion of the case was a hollow victory for the students as the court refused to enter a judgment against the state since in their opinion the State of Texas was trying to further education. The court viewed the TAKS and the multitude of opportunities to retake the test, the remedial classes to help prepare for the TAKS before and after enrolled in high school showed the state was encouraging learning.

NCLB

It was not until President George W. Bush was finalizing what is known today as the NCLB was there a major call for standardized testing. This contradicted the explicit language of preceding decades of national legislation strictly prohibiting high-stakes testing. President Bush advocated the idea of standardized testing for all students based on his success of how test scores had increased in Texas while he was Governor Bush.

It was not until Bush was well into his administration that it was revealed most of the results in Texas were flawed, if not downright criminal. An outside Texas Education Agency (TEA) had analyzed the Texas Assessment of Academic Skills (TAAS) data and found gross inaccuracies in what had been called the 'Texas Miracle'; "…passing scores set on TAAS tests were arbitrary, discriminatory and failed to take measurement error into account, analyses comparing TAAS reading, writing and math scores with one another and with relevant high school grades raise doubts about the reliability and validity of TAAS scores" (Haney, 2000, p. 120). There were also concerns raised about graduation rates calculations, deceptive enrollment policies to affect minority achievements and overall gross mismanagement of applying the results of the TAAS. However, this was the model used by Bush for the NCLB.

The NCLB was signed into law after a contentious election with presidential hopefuls Bush and Al Gore on opposite sides of the educational testing issue. Gore wanted to continue with the Democratic ideas laid out in Goals 2000 by the outgoing President Clinton of voluntary testing, while Bush fresh on the heels of a perceived turn around in Texas education wanted mandatory testing for grades 4-8. Once the chads had been counted in Palm Beach County, Florida and the courts weighing in on the election results in which Bush was declared the victor the now President Bush got to work on reauthorizing ESEA.

One of the President's first jobs is to appoint a cabinet. Bush quickly chose Rod Paige who was superintendent of Houston City Schools as his Secretary of Education. At the time, no one could argue with the apparent results of the turnaround in the Texas schools and Paige was confirmed as Secretary of Education. Bush as a result had a huge educational ally when it came to any educational issue. However, it was a relationship that Bush would not have to take advantage of due to one of the most catastrophic events in U.S. history.

The NCLB was signed after the 9/11 terrorist attacks as Bush was riding a wave of public support for his stance to search and find the perpetrators of such a calamity on the U.S. While the Democratic Senate did not agree with the mandatory testing idea of the NCLB, they knew opposition to a popular president would not sit well with their newly patriotic senatorial constituency. There was some tentativeness of his Republican counterparts. In an *Education Week* article the observation is made,

> Some conservatives are wary of the active federal role he has set out, but say they recognize the current political reality of having a Republican president who supports such a role… One of the thorniest issues for conservatives is how to demand accountability without asserting too much federal control. For the president, accountability in part means testing. He would require states to test students annually in grades 4-8, which is more frequent than most states now do. Though he has made clear that states--not the federal government--would develop those tests (Robelen, 2001).

In the Purpose statement of the NCLB, it details two references to increased accountability by assessments;

- ensuring that high-quality academic assessments, accountability systems, teacher preparation and training, curriculum, and instructional materials are aligned with challenging State academic standards so that students, teachers, parents, and administrators can measure progress against common expectations for student academic achievement;
- improving and strengthening accountability, teaching, and learning by using State assessment systems designed to ensure that students are meeting challenging State academic achievement and content standards and increasing achievement overall,

> but especially for the disadvantaged; (*An Act to Close the Achievement Gap with Accountability, Flexibility, and Choice, So That No Child Is Left Behind*, 2002, p. 115 stat. 1440).

Within the NCLB are very specific requirements for assessments of students. The NCLB outlines the general guidelines for assessments as:

> Each State plan shall demonstrate that the State educational agency, in consultation with local educational agencies, has implemented a set of high quality, yearly student academic assessments that include, at a minimum, academic assessments in mathematics, reading or language arts, and science that will be used as the primary means of determining the yearly performance of the State and of each local educational agency and school in the State in enabling all children to meet the State's challenging student academic achievement standards (*An Act to Close the Achievement Gap with Accountability, Flexibility, and Choice, So That No Child Is Left Behind*, 2002, p. 115 stat. 1449).

The NCLB then goes on to cite fourteen particular requirements that the states are to carry out in order to meet the NCLB assessment specifications, including what grades will be tested; "measure the proficiency of students in, at a minimum, mathematics and reading or language arts, and be administered not less than once during—'(aa) grades 3 through 5;'(bb) grades 6 through 9; and '(cc) grades 10 through 12;"(*An Act to Close the Achievement Gap with Accountability, Flexibility, and Choice, So That No Child Is Left Behind*, 2002, p. 115 stat. 1450). Science is to be assessed starting in 2007 according to the NCLB.

This now forced the states hand as the states were obligated to implement high-stakes testing or forego federal monies earmarked for education. Most states did not hesitate to put the federally suggested ideas into place.

Connecticut v. Spellings, 453 F. Supp. 2d 459 - Dist. Court, D. Connecticut 2006

The NCLB went relatively unchallenged for several years due to the patriotism of supporting President Bush post-911 and states trying to implement the requirements of the NCLB. It was not until 2006 that Connecticut challenged the NCLB in court. As one of the first states to challenge the NCLB, Connecticut "takes the federal government to task for failing to fund programs required by the new law, particularly in the area of state assessments. Connecticut argues that "…the education law explicitly prohibits the government from making unfunded mandates" (*Recent Court Decisions and Legislation Impacting Juveniles Recent Court Decisions*, 2006, p. 264). At issue, was the idea that the government was not allowed to pass a law that required states to abide by the NCLB without providing the funds to abide by the law, namely the high-stakes test.

The rule of law in this case was that it was unconstitutional for the government to pass an unfunded mandate. The application was the NCLB imposing this mandate that states were to follow or the federal government would withdraw education funding to that particular state. The lawsuit was eventually dismissed in 2008 based upon the states inability to prove how the NCLB had placed an unfair financial burden on the state of Connecticut. The very public case of trying to get out from under the NCLB was not in vain as only a few years later Connecticut was granted a waiver from the NCLB from the Obama administration. The waivers were issued to states in response to the NCLB's unrealistic ramped up educational goals of all children reading

on grade level by 2014. The lack of funding proved to be a minimal issue in light of the unachievable educational goals and the subsequent consequences of missing those goals.

Valenzuela v. O'Connell, No. CPF-06-506050 (San Francisco County Ct. Mar. 23, 2006)

The issue brought by the plaintiffs was that the state of California could not deny high school diplomas to students who did not pass the California High School Exit Exam (CAHSEE). The plaintiffs of this class-action lawsuit were mainly low-income, Hispanic or African-American students who were being tested as per the mandates of the NCLB.

The rule of law being tested was the harm of being denied a high school diploma. Specifically, the harms being litigated were related to equal protection and the right to an education. The plaintiffs were arguing that by denying a high school diploma to students that had otherwise been eligible to receive a diploma was suffering undo harm by the newly implemented CAHSEE.

The trial court judged agreed with the plaintiffs and applied the rule of law to this case by enjoining the district from denying the plaintiffs a diploma. While the plaintiffs were given an extra two years to pass the exit exam as a recognition of a new graduation requirement being unfair to some, this additional time was not helpful to these students. Students were not passing the exam after these two extra years and the state of California was prepared to deny thousands of students in California their high school diploma only due to those students not passing the CAHSEE.

This victory was short-lived as an appeals court overturned the lower court decision due to the appeals court view that this was an "improper encroachment onto legislative terrain" (Heise, 2009, p. 6). This avoidance of ruling on educational legislation resounded throughout the legal landscape. This ruling allowed jurists to wave off any rulings that would contradict the

drafted legislation. This gave the educational policymakers wider latitude when designing educational legislation knowing the courts were resistant to affect many types of educational legislation. Heise (2009a) surmises the reasoning behind the courts actions for ruling for the defendants was to avoid the superseding of legislation that would "involve efforts to shore up the currency of the high school diploma and to improve student and school performance" (p. 7).

Race to the Top

The fallout from the NCLB led to many states to eye 2014 that all students are on grade level in reading and mathematics as a deadline not a finish line. This unachievable goal led to two items 1) the Obama administration sponsoring another federal incentive program called, Race to the Top (RTTT) and 2) states petitioning the federal government for waivers from the NCLB. Both had effects on high-stakes testing.

RTTT involved states competing for federal monies by:

> The ARRA [American Recovery and Reinvestment Act of 2009] provides $4.35 billion for the Race to the Top Fund, a competitive grant program designed to encourage and reward States that are creating the conditions for education innovation and reform; achieving significant improvement in student outcomes, including making substantial gains in student achievement, closing achievement gaps, improving high school graduation rates, and ensuring student preparation for success in college and careers; and implementing ambitious plans in four core education reform areas (U. S. Department of Education, 2009, p. 2).

This race was hard for many states to ignore due to the perceived extra monies to be funneled to education. RTTT does specifically address testing, "(B)(2) Developing and implementing common, high-quality assessments *(10 points)* (B)(3) Supporting the transition to

enhanced standards and high-quality assessments *(20 points)*" (U. S. Department of Education, 2009, p. 3). However, only 30 points out of a total of 485 points were dedicated to testing. This equates to 6 % of the total points was dedicated to an already flawed assessment system. It makes one wonder what other areas of RTTT are also flawed. "The U.S. Department of Education via congressional authorization is violating state and local rights to control education reform" (Barnes, 2011, p. 397). This shows that despite the multitude of points and incentives to improve education, the federal government is overstepping its governance limits.

Waivers

> When the framers of the United States Constitution wrote that it is a duty of the chief executive to 'take care' that the laws be faithfully executed, they can hardly have imagined a law so freighted with perverse and destructive consequences as No Child Left Behind (Derthick & Rotherham, 2012, p. 57).

This shows that a little forethought by the Bush Administration may have alleviated some pain and agony that students and educational stakeholders had to endure since the passage of the NCLB. Testing is but one destructive component among the myriad of deficiencies of the NCLB. The Obama administration has recognized some of the shortcomings of the NCLB and is trying to make provisions to fix the NCLB, mainly by issuing waivers.

The idea behind the NCLB waivers is to allow states to petition the federal government to no longer be subjected to the NCLB provisions while still maintaining a promise to improve education within the state. However, this has been met with criticism:

> Where Secretary Duncan's waivers get complicated is the hodgepodge of laws, regulations, and initiatives that comprise federal education policy today, again because of congressional inaction. The federal goals of improving teacher evaluations, adopting college- and career-ready standards, and turning around low-performing schools trace their legislative provenance to congressional authorizations permitting the secretary of education to allocate federal funds based on priorities he determines rather than specific laws passed by Congress (Derthick & Rotherham, 2012, p. 60).

Rotherham (2012) points out that while the NCLB waivers are a good idea the implementation is yet another mistake.

> Allowing an arbitrary federal appointee like Education Secretary Arne Duncan to determine if a state should be allowed to veer from a federal mandate based on subjective analysis is on tenuous legal ground to say the least (page 61).

Again, the sufferers are the students. It is the students that still must take some type of high-stakes test to maintain the appearance of accountability. The Obama administration has suggested moving away from high-stakes testing but they have no plan on how to hold schools accountable without testing.

CHAPTER 3

ANALYSIS OF THE LEGAL HISTORY AND LEGISLATION

Chapter 3 presents an analysis of the legal history and legislation that affected high-stakes testing to answer the research questions:

1. What is the relevant legal history of high-stakes testing?
2. What are the predominate issues that have been litigated in the courts relevant to high-stakes testing?
3. What is the current legal status of high-stakes testing?

High-stakes testing is a type of education reform borne from state and later federal legislative action. These legislative actions have been challenged in the courts and many related issues have been addressed by courts. Chapter 2 outlined the chronological history of high-stakes legislation and the seminal court cases.

Table 3.1

High-stakes testing related court cases with the constitutional issues and final court decisions

Case	**Constitutional Issue**	**Final Court Decision in favor of**
Brown v. Board of Education	Equal Protection	Plaintiff
San Antonio v. Rodriguez	Equal Protection	Defendant
Scheelhaase v. Woodbury Sch.	Due Process	Defendant
Debra P. v. Turlington	Due Process	Plaintiff
Rankins v. LA State Bd. Ed.	Equal Protection	Defendant
Erik V. v. Causby	Equal Protection	Defendant

GI Forum v. Tex. Educ. Ag.	Due Process	Defendant
Connecticut v. Spellings	Unfunded Mandate*	Defendant
Valenzuela v. O'Connell	Equal Protection	Defendant

*Unfunded mandate is not a national constitutional issue

Predominate Issues and Current Legal Status

Equal Protection

Five of the ten cases analyzed were brought to court citing the equal protection provision of the U.S. Constitution. *Brown v. Board of Education* was the only case where the plaintiffs were victorious.

> We conclude that in the field of public education the doctrine of 'separate but equal' has no place. Separate educational facilities are inherently unequal. Therefore, we hold that the plaintiffs and others similarly situated for whom the actions have been brought are, by reason of the segregation complained of, deprived of the equal protection of the laws guaranteed by the Fourteenth Amendment. This disposition makes unnecessary any discussion whether such segregation also violates the Due Process Clause of the Fourteenth Amendment ("Brown v. Board of Education," 1954).

The effects of this one victory forever changed educational policies in the U.S. This case paved the way for many more education cases to be filed under the equal protection clause. The historic implications of this decision were felt across the U.S. as all school districts now had to provide equal access for all students without regard to race.

Most high-stakes testing cases filed cited the equal protection clause. Since *Brown v. Board*, decided under the equal protection clause, many plaintiffs used that precedent for a

successful judgment by arguing their case violated equal protection rights. These cases generally referenced the *Brown* decision to strengthen their arguments.

San Antonio v. Rodriguez prominently cited the *Brown* case, however the court struck down the argument that equal educational funding is a fundamental right under the U.S. Constitution. The Court concluded:

> The consideration and initiation of fundamental reforms with respect to state taxation and education are matters reserved for the legislative processes of the various States, and we do no violence to the values of federalism and separation of powers by staying our hand. We hardly need add that this Court's action today is not to be viewed as placing its judicial imprimatur on the status quo. The need is apparent for reform in tax systems which may well have relied too long and too heavily on the local property tax. And certainly innovative thinking as to public education, its methods, and its funding is necessary to assure both a higher level of quality and greater uniformity of opportunity. These matters merit the continued attention of the scholars who already have contributed much by their challenges. But the ultimate solutions must come from the lawmakers and from the democratic pressures of those who elect them ("San Antonio Independent School Dist. v. Rodriguez," 1973).

This shows that the majority of the Court was apprehensive in adjudicating educational issues related to economic rather than racial inequalities despite the insistence of the plaintiffs' claims of violation of constitutional equal protection rights:

> Thus, we stand on familiar ground when we continue to acknowledge that the Justices of this Court lack both the expertise and the familiarity with local problems so necessary to

> the making of wise decisions with respect to the raising and disposition of public revenues ("San Antonio Independent School Dist. v. Rodriguez," 1973).

One can see why this case was decided by the slimmest of margins, 5-4, in favor of reversing the lower court decision. The minority opinion did address the equal protection argument in this case:

> Requiring the State to establish only that unequal treatment is in furtherance of a permissible goal, without also requiring the State to show that the means chosen to effectuate that goal are rationally related to its achievement, makes equal protection analysis no more than an empty gesture… the parents and children in Edgewood, and in like districts, suffer from an invidious discrimination violative of the Equal Protection Clause ("San Antonio Independent School Dist. v. Rodriguez," 1973).

These dissenting justices decided that equal protection should have been the main issue. They felt that constitutional rights should have taken precedence over state or local jurisdictional policies or laws. Justice Marshall went on at length and voiced his displeasure with the majority opinion and summarized his discontent with the decision and how the case violated the equal protection clause:

> The Court seeks solace for its action today in the possibility of legislative reform. The Court's suggestions of legislative redress and experimentation will doubtless be of great comfort to the schoolchildren of Texas' disadvantaged districts, but considering the vested interests of wealthy school districts in the preservation of the status quo, they are worth little more. The possibility of legislative action is, in all events, no answer to this Court's duty under the Constitution to eliminate unjustified state discrimination. In this case we have been presented with an instance of such discrimination, in a particularly invidious

> form, against an individual interest of large constitutional and practical importance. To support the demonstrated discrimination in the provision of educational opportunity the State has offered a justification which, on analysis, takes on at best an ephemeral character. Thus, I believe that the wide disparities in taxable district property wealth inherent in the local property tax element of the Texas financing scheme render that scheme violative of the Equal Protection Clause ("San Antonio Independent School Dist. v. Rodriguez," 1973).

While the majority opinion and the final adjudication of this case were a loss for the plaintiffs, this case continued to inspire challenges to inequities arguing that equal protection can be a right violated by educational policies that result in funding disparities. The closeness of this decision gave hope to latter plaintiffs that in states courts, and possibly federal courts, there were judges that would hold that constitutional rights trump politics in educational policies.

Rankins and *Valenzuela* represent cases where constitutionality was set aside and the conflict of court rulings on educational policy took priority. The court felt that these cases were based on legislative issues and not constitutional ones.

Rankin's conclusion by the Louisiana Appellate Court references the *Debra P*. case and makes clear the court's hesitancy to cross jurisdictional boundaries due to a state educational goal:

> Although the state provides certain funds for services which directly assist students in non-public schools, we join the *Debra P*. court in acknowledging that the state has a stronger interest in the public education for which it pays the costs. We conclude that the GEE does not violate the equal protection clause because its administration is rationally related to the state's legitimate interest of insuring minimum competency among persons

obtaining a state diploma ("Rankins v. Louisiana State Bd. of Elementary & Secondary Educ.," 1994).

The *Valenzuela* case was decided by a judgment in the Court of Appeals of California. "…the trial court erred in granting a statewide preliminary injunction enjoining defendants from enforcing the statute mandating the CAHSEE diploma" ("O'Connell v. Superior Court," 2006). The exit exam requirements for the state of California were deemed proper according to the appellate court because:

> In 1999, the Legislature decided that one way to address the inadequacy of California's education system was to create the CAHSEE, and to require students to pass it in order to receive their high school diplomas, while providing remedial education for those students not yet having the skills to pass. The Legislature also conferred discretion on the executive branch to determine that the CAHSEE diploma requirement would apply to the high school class of 2006, and that no alternatives would be adopted. Those actions are entitled to substantial deference by the judicial branch, which is constitutionally obligated to refrain from usurping the role of the other two branches in formulating and implementing public policy ("O'Connell v. Superior Court," 2006).

This points out that the Appeals Court of California did not want to overstep their judicial boundaries. However, the court did concur with the lower court ruling on equal protection:

> Given the standard of review and our normal deference to trial court findings of fact, we accept the trial court's conclusion that plaintiffs established a likelihood of success on the merits as to the denial of their fundamental right to equal educational opportunity ("O'Connell v. Superior Court," 2006).

The issue of equal protection was disregarded as the court ruled in favor of the defendant. It does seem odd that while the court agreed with the merits of the argument of equal protection the court did not address the issue in their final decision. The choice to not address equal protection in the final decision could be interpreted as a victory for supporters of equal protection with respect to high-stakes testing. Since the court did not specifically rule against the plaintiffs' charges that a CAHSEE diploma should fall under the equal protection provision of the Constitution, it allowed high-stakes testing plaintiffs the opportunity to continue to argue for high-stakes testing denial of diplomas to be considered under the equal protection clause.

The *Eric V. v. Causby* case closely mirrors the *Debra P. v. Turlington* case, which will be discussed under the Due Process part of this chapter, but *Eric V*. does not reach the same type of notoriety because the court struck down the heart of the plaintiffs' argument. It did take seriously the equal protection argument when weighing the facts of the case however:

> A county school board has an indisputable interest in implementing its own policy regarding promotion of students based on qualitative achievement standards. A 'classification' based on students' scores on standardized tests is surely the paradigmatic situation for application of rational basis review… the Johnston County Board of Education has chosen a rational means, the end-of-grade test, to foster a legitimate end, encouragement of academic achievement. These conclusions are, of course, subject to being disproved on the facts, but Plaintiffs, at this stage, have failed to prove a high likelihood of success on such argument ("Erik V. , by and through Catherine V. v. Causby," 1997).

The result shows that a lower court can weaken a plaintiffs' case by deciding the constitutional arguments are not valid. While *Brown* may have energized potential plaintiffs to

sue based upon constitutional equal protection rights, cases like *Erik V. v. Causby* indicated it can be an arduous path for courts to rule in favor of the plaintiff based solely on this constitutional right. Since the *Eric V.* case was settled in 1997, it is seems the importance of a monumental decision like *Brown* is carrying less weight.

Due Process

The three due process cases identified were decided two for the defendants and one for the plaintiff. The *Debra P. v. Turlington* case was a landmark decision because the justices felt some type of implementation guidelines needed to exist. The justices felt that Debra P. did not get enough notice of the state's high-stakes test and that was in violation of her constitutional due process rights.

> Pursuant to the findings in Sections IV D. and V, the Court is of the opinion that declaratory and injunctive reliefs are both appropriate and proper in the present instance. In a separate Order the Court will declare that Fla.Stat. § 232.246(1)(b) (1978 Supp.) is, as applied, in the present context a violation of the equal protection and due process clauses of the Fourteenth Amendment. 42 U.S.C. § 2000d, and 20 U.S.C. § 1703 ("Debra P. v. Turlington," 1979).

The ruling made clear that there needed to be a legally appropriate time of notice and preparation available before implementing these tests. While the due process aspect of this case was a victory, on appeal other parts of the case were struck down:

> We affirm the district court's findings (1) that students were actually taught test skills, (2) that vestiges of past intentional segregation do not cause the SSAT-II's disproportionate impact on blacks, and (3) that use of the SSAT-II as a diploma sanction will help remedy the vestiges of past segregation. Therefore, the State of Florida may deny diplomas to

students (beginning with the Class of 1983) who have not yet passed the SSAT-II ("Debra P. By Irene P. v. Turlington," 1984).

The *Scheelhaase* decision was more of a workplace dispute stemming from the dismissal of a teacher. High-stakes testing did play a role since her dismissal was based on her inability to make students successful on the state-mandated tests. “The specific reason given plaintiff for termination was her professional incompetence as indicated by the low scholastic accomplishment of her students on the Iowa Tests of Basic Skills (ITBS) and Iowa Tests of Educational Development (ITED)” ("Scheelhaase v. Woodbury Central Community Sch. Dist," 1972). The plaintiff felt there was a violation of her due process rights which the court did agree with, “Due process demands that reasons for termination of a teacher's contract may not be arbitrary and capricious but must have a basis in fact” ("Scheelhaase v. Woodbury Central Community Sch. Dist," 1972). Upon appeal these issues were struck down.

> The Superintendent and the Board for the Woodbury, Iowa, Central Community School District possessed the right and responsibility of evaluating its teacher personnel, and such evaluations, where they are based on some evidence, even though possibly erroneous, will not serve to make those determinations subject to judicial review as unconstitutionally arbitrary and capricious ("Scheelhaase v. Woodbury Central Community Sch. Dist," 1973).

The *GI Forum Image De Tejas v. Tex. Educ. Agency (TEA)* decision opined due process was met:

> The Court concludes that the TAAS test violates neither the procedural nor the substantive due process rights of the Plaintiffs. The TEA has provided adequate notice of the consequences of the exam and has ensured that the exam is strongly correlated to

material actually taught in the classroom ("GI Forum Image De Tejas v. Texas Educ. Agency," 2000).

The U.S. District court did establish a solid benchmark on the issue of due process in this case, however they also affirmed their reluctance to rule on jurisdiction issues, writing: “It is not for this Court to determine whether Texas has chosen the best of all possible means for achieving these goals. The system is not perfect, but the Court cannot say that it is unconstitutional” ("GI Forum Image De Tejas v. Texas Educ. Agency," 2000). This can be viewed as weakening their opinion on due process since the Court did not want to foray into areas it deemed inappropriate.

The *Debra P.* due process type of case was more common when high-stakes tests were first being given. According to cases that were filed, many states were not giving either enough notice of a high-stakes test or preparation for the high-stakes test. These issues have since been addressed to the appeasement of the courts. Since the NCLB made all states carry out the testing mandates, states learned how to abide by due process standards. With respect to the *GI Forum* and *Scheelhaase* cases, once due process was viewed as a non-issue, the case was difficult to move forward to the plaintiffs’ liking.

Unfunded Mandate

The original complaint brought to the U.S. District Court of *Connecticut v. Spellings* of great significance to both supporters and opponents of high-stakes testing. The focus of this case was whether the NCLB could be forced upon the states without federal funding.

> In this case, the State challenges the Secretary's interpretation of several key elements of the Act. In particular, the State asks the Court to clarify the meaning of the so-called ‘Unfunded Mandates Provision’ of the Act, 20 U.S.C. § 7907(a), and to declare that the Secretary's interpretation of that provision is contrary to its plain language and Congress's

> intent in enacting it. In addition, the State seeks a ruling that the Secretary's implementation of the Act violates both the Spending Clause of the United States Constitution and the Tenth Amendment. Finally, the State alleges that the Secretary violated the Administrative Procedures Act (the "APA"), 5 U.S.C. §§ 701-706, by denying the State's requests for waivers from the Act's requirements and also by denying certain plan amendments submitted by the State ("Connecticut v. Spellings," 2006).

This case examined many jurisdictional issues and interpretations of state and federal statutes. Connecticut had to amend the original complaint due to the decision of the judge:

> The Court GRANTS IN PART and DENIES IN PART the Secretary's Motion to Dismiss [doc. # 18]. The Motion to Dismiss is GRANTED as to Count I, Count II, and Count III in their entirety, as the Court concludes that it lacks subject matter jurisdiction over those claims. The motion is also GRANTED as to the portion of Count IV that alleges that the Secretary failed to provide the hearing required by the Act, as the Court concludes that this claim is moot. The Motion to Dismiss is DENIED as to the remaining claims in Count IV. As a result of this ruling, Count IV is the only remaining count of the State's Complaint. The parties shall file no later than October 16, 2006, a joint report setting forth their proposed schedule(s) for filing the administrative record and for resolving the remaining claims in this action. Having now determined what claims are before it, the Court orders that any non-party wishing to intervene on the issues in Count IV must file a motion to intervene and supporting papers no later than October 16, 2006 ("Connecticut v. Spellings," 2006).

The state re-filed as instructed but failed to impress the judge:

> The Court wishes to be clear that it has not ruled on the merits of the State's Unfunded Mandates Provision claim because the argument was never made in connection with these two proposed plan amendments. Therefore, the State is free to pursue that issue before the Secretary. It is truly unfortunate that the Court is unable to reach this issue because the State failed adequately to raise it in the context of the State's proposed plan amendments. For immediately after the Court's ruling on the Secretary's Motion to Dismiss, the Court suggested to the State that it consider dismissing Count IV without prejudice in order to allow the State to return to the Secretary to develop a detailed record regarding the State's unfunded mandates argument. Instead, the State decided to continue to litigate the issue in this Court. Regrettably, the result is that over a year and one-half after the Motion to Dismiss Ruling, the State is no closer to a determination of this very important issue ("Connecticut v. Spellings," 2008).

Connecticut did not re-file their complaint possibly due to a new presidential administration in 2008 who appointed a new Secretary of Education. The unfunded mandate issue essentially died once President Obama and Secretary of Education Arne Duncan announced the 2011 waivers for states from the NCLB. With the issuance of waivers to most states requesting waivers it allows states to oversee education goals without the punitive consequences of the federal mandate of the NCLB.

Currently the challenges to the legal status of high-stakes testing are coming from grass-roots campaigns to repeal or eliminate high-stakes tests. The fairtest.org website contains a statement addressing the multitude of parents, students, teachers and administrators who are unhappy with the effects of high-stakes testing:

> A nationwide protest movement against the stranglehold of high-stakes testing on our schools has escalated to a rolling boil. Boycotts, opt-out campaigns, demonstrations, and community forums are among the tactics being pursued in cities such as Austin, Seattle, Portland, Oregon, Chicago, Denver and Providence. Meanwhile, the number of signers of the National Resolution on High-Stakes Testing continues to grow ("Test Opposition Surges Across the Nation," 2013).

These types of challenges to educational policy or legislation have affected the ESEA re-authorization. The most recent legislative authors of a re-authorization bill seem to have realized the issues educational stakeholder had problems with in the NCLB. Senator Tom Harkin, a Democratic senator from Iowa, and Chairman of the Senate Health, Education, Labor, and Pensions (HELP) Committee, introduced in June of 2013 the Strengthening America's Schools Act of 2013 which states:

> If a state has an accountability system approved by the Secretary, it can continue to use their approved accountability system. If not, a state will adopt an accountability system that is equally ambitious and holds all students to high expectations of student achievement.
>
> All accountability systems will include student academic achievement and growth, English language proficiency for English Learners and, for high schools, graduation rates for all students; systems will also include accountability for all subgroups. This accountability system asks states to identify and support –
>
> Priority schools - The lowest-achieving 5 percent of each elementary schools and secondary schools, and secondary schools with a graduation rate lower than 60 percent.

Focus Schools - Ten percent of schools with the greatest achievement gaps and secondary schools with the greatest graduation rate gaps between subgroups.

For all other schools, districts will identify schools experiencing achievement gaps across subgroups and will develop and implement a locally-designed intervention for that school based on input from the community.

In conversations with Congressional representatives for Georgia, Senator Isakson, Senator Chambliss and Representative Woodall they all felt that because of party divisions the bill does not stand much chance of passing or receiving a presidential signature. In fact, there are currently three bills in Congress to re-authorize ESEA, however bipartisan politics is expected to keep them from passing.

These three bills do share the issue of testing and data being used to evaluate students. Local educational stakeholders and the federal government have divergent ideas on how to measure education. High-stakes testing is still supported by the federal government, while at the local level stakeholders are voicing the need to eliminate or curtail high-stakes testing. It appears there is a growing divide between what the local entities want and the laws Congress is drawing up.

Most of the constitutional issues related to the various components of high-stakes testing have been heard. However, as educational reform continues, high-stakes testing is sure to create new constitutional challenges in the future. The influence local stakeholders have seems to be on the rise as well. Where under the NCLB it was the federal government that held the authority it is now those affected by the NCLB who are fighting back against that authority.

CHAPTER 4

SUMMARY, FINDINGS AND CONCLUSIONS

The purpose of this study was to review and analyze important historical and legal events related to high stakes testing and its impacts on students. This was accomplished by analyzing relevant legislation and landmark court cases that dealt with high-stakes testing.

Federal policymakers were writing legislation in reaction to conditions set forth in federal legislation such as ESEA or national reports such as A Nation at Risk. Local policymakers struggled with being constitutionally compliant when implementing high-stakes tests such as *Debra P v. Turlington*. State and local courts tested education laws, policies, and their implementation on issues such as due process and equal protection in deciding high-stakes testing cases.

The legislation reviewed was precedent setting education legislation initiated by a presidential administration or Congress.

The landmark cases started with the seminal equal protection and discrimination case of *Brown v. Board.* This case changed the educational landscape for decades to come. As a result of *Brown v. Board* many high-stakes testing court cases and legislation were borne from the impact of preserving constitutional rights for all students regardless of skin color or any other categorizing stigma. Based on the decision of the Court on *Brown*, other courts were careful to abide by the tenets set forth by the *Brown* case. This was seen in *San Antonio v. Rodriguez* which cited *Brown* as a basis of the plaintiffs' argument.

Findings and Conclusions

Based on a review of relevant history and law, this study finds and concludes the following:

Circumventing Research Suggestions

Despite salient, research-based findings of the inherent unfairness of high-stakes testing, legislation continued to be written and implemented expanding high stakes testing. The top research associations in the United States published the *Standards for Educational Research and Testing* five different times since 1954. Yet much of the criteria required have been seemingly ignored. The committee for *Standards* defined high-stakes testing and enumerated specific guidelines to follow to implement a fair test. While there were some references in legislation and court cases, the *Standards* seem to have been largely ignored.

The four categories of fairness and twelve standards of fairness are non-binding legally of course, yet most experts would agree that they are a keystone to designing a fair test.

Fairness as a lack of bias

Fairness as a lack of bias was violated in numerous cases. As an example, in *Brown v. Board* the court found that black schoolchildren were subjected to sub-standard education. The fairness standard states bias occurs "when deficiencies in a test itself of the manner in which it is used result in different meanings for scores earned by members of different identifiable subgroups" (American Educational Research Association et al., 1999). While this case is not about a high-stakes test, to forbid children of color to attend school, or assign the grossly inequitable resources, was a bias against groups of schoolchildren on the basis of color alone. High-stakes tests violate this bias if questions are culturally biased and near impossible for students outside of that particular cultural upbringing to answer correctly. According to the

website fairtest.org, “Biased cultural assumptions built into the test as a whole often are not removed by test-makers” ("What's Wrong With Standardized Tests?," 2012).

Fairness as Equitable Treatment in the Testing Process

Several court cases ruling against high-stakes testing pointed out this standard was violated. As the standard states, “Fair treatment of all examinees require consideration not only of a test itself, but also the context and purpose of testing and the manner in which test scores are used” (American Educational Research Association et al., 1999). In *Scheelhaase* this was violated by using the test scores against a teacher. The scores were supposed to be used to show some type of academic mastery, but instead were used as evidence to show that the teacher was not educating her students. These tests were not intended to set a basis of whether to keep or fire the teachers. *GI Forum* challenged the use of high-stakes testing as a benchmark to receive a high school diploma. The plaintiffs asserted the test unfairly targeted minorities. Even though the court ruled against the plaintiffs, the results of the test could not be disputed as minorities continued to fail the high-stakes test. This could be viewed as using a test for questionable purposes.

Fairness as Equality in the Outcomes of Testing

Arguably the biggest offender of this standard is the NCLB. Many of the consequences of the NCLB were not based on valid evidence nor were the interpretations of the wide spread high-stakes test. The evidence was supposed to have been Bush’s success in Texas with testing. However, it was later proven that there were gross mismanagement of implementation of the tests and interpretation of the results. (Haney 2000). The validity of the intention to use a standardized test to measure all students learning, regardless of their mental capabilities or deficiencies is non-existent. The idea that mentally challenged students would be reading on the

same grade level by 2014 is not only unachievable, but irresponsible of any legislator to draft or support such legislation. The equality of outcome is problematic as states, districts, teachers and students are judged by the outcomes of these invalid tests and upon failing, would be branded as failing and subject to closure. To categorize mentally challenged students the same as mentally capable students is unrealistic, resulting in negligent legislation and policy at best.

Fairness as Opportunity to Learn

The seminal case of *Debra P. v. Turlington* displays a clear violation of this standard. Debra P. failed the high-stakes test and was denied a high school diploma. Since according to the standard, "When test takers have not had the opportunity to learn the material…" which Debra P. did not, this case unmistakably shows how this test was unfair. The court ruled that Debra P.'s property rights to a diploma had been violated by the state of Florida by not allowing adequate time for students like Debra P. to learn the material. While many states have since learned from this case, one can question whether the withholding of a diploma based upon the failure of one test is fair.

Fairness and the Constitution

A fundamental principle underlying the Constitution is fairness to all citizens (e.g., due process; equal protection). "Substantive due process requires that government actions must be fundamentally fair" (Dayton, 2012, p. 241). In complex terms equal protection according to the Court is trickier as Dayton (2012) explains there are three levels of scrutiny the Court uses to decide cases claiming equal protection. The simpler terms of equal protection Dayton (2012) observes is, "Fundamental fairness requires equal treatment of all persons in equal circumstances (Dayton, 2012, p. 267).

This study reviewed constitutional principles and research findings that reflect the mirror of the fairness standards set forth in the Constitution and by the AERA committee. Courts sometimes found high-stakes testing in violation of constitutional principles. But it can be argued that these same high-stakes tests violate this fairness idea even when the courts are either hesitant or unable to rule on fundamental fairness issues.

Avoiding adhering to the *Standards* by legislators might be reason enough to question the validity of high-stakes testing. However, Congress under the Clinton administration received a report on the foibles of high-stakes testing by this very group. Congressional questioning of testing produced several recommendations with some intriguing concepts, "…High-stakes decisions …should not be made on the basis of a single test score" (Heubert et al., 1999). Current legislation, such as the NCLB, seems to have ignored its own Congressional research. Congress also ignored this recommendation: "Policymakers should monitor both the intended and unintended consequences of high-stakes assessments" (Heubert et al., 1999). Congress was supposed to revisit the NCLB in 2007, yet the reauthorization has not happened as of this writing in 2013. If Congress is unwilling to analyze this legislation it leaves those disadvantaged by such legislation with little recourse, since courts tend to be highly deferential to Congress, ultimately side with the defendants in most high-stakes testing cases.

Due process challenges have diminished since the mandatory high-stakes tests implementation deadline of 2006 by the NCLB. Since due process includes allowing a fair time period for the implementation most of those challenges would be moot today since the NCLB was passed over a decade ago.

Equal protection is a valid argument on its face, but the courts have been hesitant to rule on students as a class of people being constitutionally harmed since the *Brown* decision. Most

courts feel that students come under the guise of local educational authorities. Since, according to the courts, these authorities are better versed on what is best for students.

> Requiring the State to establish only that unequal treatment is in furtherance of a permissible goal, without also requiring the State to show that the means chosen to effectuate that goal are rationally related to its achievement, makes equal protection analysis no more than an empty gesture ("San Antonio Independent School Dist. v. Rodriguez," 1973).

Reign of Error

The Ravitch book, *Reign of Error* points out the fallacies of high-stakes testing reforms while offering solutions to counteract these well intentioned yet wayward reforms. The book gives hope to the idea of change, since Ravitch herself changed her viewpoint on high-stakes tests. As a Washington insider and one who suggested and implemented policy many can take hope that others inside Washington will have a change of heart towards this type of testing that Ravitch did. What Ravitch offers are pragmatic solutions to what has become a complicated issue. Ravitch ignores the legal issues of fairness and focuses on the history of education. Her book offers solutions based on the history of what students need to learn to be productive citizens. *Reign of Error* is a reminder of the basic tenets of education and illustrates how far education has strayed from those tenets under the guise of reform.

Conclusions

This study found that legislators, policymakers, and the courts have largely ignored *Standards for Educational Research and Testing* and other research on high-stakes testing. Commonly high-stakes testing has been enabled by legislation that violated basic constitutional rights such as due process and equal protection.

Legislation that violates basic constitutional rights should not be implemented because it is unconstitutional, but even more so when it is also bad policy. Educational legislation establishing the current high-stakes testing regime ignores sound research and history. Some high-stakes tests are known to have culture and racial bias such as the Texas TAKS test as analyzed by fairtest.org article *Racial Bias Built into Tests* and the IQ tests in the case *Larry P v. Riles*. These are similar to issues argued in the *Brown* proceedings, yet the further we get from the *Brown* case, the less we seem to have learned.

With all the scientific research admonishing high-stakes testing the courts have yet to realize the magnitude of denying student requests for relief based on federal or local constitutional violations. Students are becoming disenfranchised with the schools that are supposed to engage them in society. It is easy to understand why students are taking matters into their own hands by refusing to take high-stakes tests. (Dornfield, 2013).

There needs to be a dialogue between all stakeholders focusing on what is best for students. Most of the high-stakes legislation takes student success as a goal with a political agenda attached. It is up to the parents, students and educators to take control of their children's education. Legislators should be implementing what these local educational stakeholders want and what is based on sound research. The true high-stakes test question is whether the years of damage done by past high-stakes tests can be rectified before these educational victims are asked to lead our country, sit on our court benches and lead students in the classroom.

References

An Act to Close the Achievement Gap with Accountability, Flexibility, and Choice, So That No Child Is Left Behind. (2002). [Washington, D.C. : U.S. G.P.O. : Supt. of Docs., U.S. G.P.O., distributor, 2002].

American Educational Research Association, Washington D. C., American Psychological Association, Washington D. C., & National Council on Measurement in Education, Washington D. C. (1999). *Standards for Educational and Psychological Testing*.

Baines, Lawrence A., & Stanley, Gregory Kent. (2004). High-Stakes Hustle: Public Schools and the New Billion Dollar Accountability. *Educational Forum, The, 69*(1), 8-15.

Barnes, Catherine R. (2011). "Race to the Top" Only Benefits Big Government. *Journal of Law & Education, 40*(2), 393-402.

Brown v. Board of Education, No. No. 1, 347 483 (Supreme Court 1954).

Camara, Wayne , & Ernesto, Marianne. (2011). Updated Standards for Educational and Psychological Testing is released for public comment. Retrieved June 6, 2013, from http://www.apa.org/science/about/psa/2011/01/testing.aspx#watson

Carnoy, M, Elmore, R, & Siskin, LS. (2003). The new accountability. *High schools and highstakes testing. New York: RoutledgeFalmer*.

Childs, Ruth Axman, Eric Clearinghouse on Tests, Measurement, Evaluation, Washington D. C., & American Institutes for Research, Washington D. C. (1990). *Legal Issues in Testing*.

Christie, Samuel G., & Casey, John A. (1983). Heading Off Legal Challenges to Local Minimum Competency Programs. *Educational Evaluation & Policy Analysis, 5*(1), 31-42.

Code of Fair Testing Practices in Education (Revised). (2005). *Educational Measurement: Issues & Practice, 24*(1), 24.

Congress, US. (1994). Goals 2000: Educate America Act. *Public Law*, 103-227.

Connecticut v. Spellings, No. No. 3: 05CV1330 (MRK), 453 459 (Dist. Court, D. Connecticut 2006).

Connecticut v. Spellings, No. No. 3: 05CV1330 (MRK), 549 161 (Dist. Court, D. Connecticut 2008).

Crisafulli, Timothy P. (2006). No EDUCATOR LEFT UNSCATHED: HOW NO CHILD LEFT BEHIND THREATENS EDUCATORS' CAREERS. *Brigham Young University Education & Law Journal*(2), 613-637.

Dayton, John. (2012). *Education Law:Principles, Policies and Practice*: Wisdom Builders Press, Bangor, Maine.

Dayton, J. (2013). *Georgia Education Law A State Law Companion to John Dayton's Education Law: Principles, Policies and Practice* (First ed.). Bangor, Maine: Wisdom Builders Press.

Debra P. By Irene P. v. Turlington, No. No. 83-3326, 730 1405 (Court of Appeals, 11th Circuit 1984).

Debra P. v. Turlington, No. No. 78-892 Civ. TC, 474 244 (Dist. Court, MD Florida 1979).

DeBray, Elizabeth H. (2006). *Politics, ideology, and education*: Teachers College Press.

Department of Education, U.S. (2009). Race to the Top Executive Summary. Retrieved July 16, 2012, from www2.ed.gov/programs/racetothetop/executive-summary.pdf

Department of Education, Washington D. C. (1991). *America 2000: An Education Strategy*.

Derthick, Martha, & Rotherham, Andy. (2012). Obama's NCLB Waivers. *Education Next, 12*(2), 56-61.

Dornfield, Ann. (2013). Seattle High School's Teachers Toss District's Test: National Public Radio, Inc.

Erik V. , by and through Catherine V. v. Causby, No. No. 5: 97-CV-587-BO (2), 977 384 (Dist. Court, North Carolina 1997).

Fuhrman, Susan H., & Consortium for Policy Research in Education, New Brunswick N. J. (1993). *Designing Coherent Education Policy: Improving the System*.

Gerber, Betsy A. (2002). HIGH STAKES TESTING: A POTENTIALLY DISCRIMINATORY PRACTICE WITH DIMINISHING LEGAL RELIEF FOR STUDENTS AT RISK. *TEMPLE LAW REVIEW, 75*, 863-890.

GI Forum Image De Tejas v. Texas Educ. Agency, No. No. Civ. A. SA-97-CA1278-EP, 87 667 (Dist. Court, WD Texas 2000).

Goals 2000 educate America : the national education goals. (1993). [Washington, D.C.? : U.S. Dept. of Education, 1993?].

Haertel, Edward H., & Herman, Joan L. (2005). A Historical Perspective on Validity Arguments for Accountability Testing. *Yearbook of the National Society for the Study of Education, 104*(2), 1-34.

Haertel, Edward H., & Lorie, William A. (2004). Validating Standards-Based Test Score Interpretations. *Measurement: Interdisciplinary Research and Perspectives, 2*(2), 61-103.

Haney, Walt. (2000). The Myth of the Texas Miracle in Education. *Education Policy Analysis Archives, 8*(41).

Haney, Walt, & Madaus, George. (1978). Making Sense of the Competency Testing Movement (Vol. 48, pp. 462-484).

Heise, Michael. (2009). COURTING TROUBLE: LITIGATION, HIGH-STAKES TESTING, AND EDUCATION POLICY. *Indiana Law Review, 42*(2), 327-342.

Heubert, Jay P., Hauser, Robert M., & National Academy of Sciences - National Research Council, Washington D. C. (1999). *High Stakes: Testing for Tracking, Promotion, and Graduation*.

Jennings, John F. (1998). *Why national standards and tests? : politics and the quest for better schools / by John F. Jennings*: Thousand Oaks, Calif. : Sage Publications, c1998.

Johanningmeier, Erwin V. (2010). "A Nation at Risk" and "Sputnik": Compared and Reconsidered. *American Educational History Journal, 37*(2), 347-365.

Johnson, E. Lloyd. (1995). THE LOUISIANA GRADUATION EXIT EXAM: PERMISSIBLE DISCRIMINATION. *Southern University Law Review, 22*, 185.

Koretz, D. (2002). High-stakes testing—where we've been and where we are. *Retrieved July*, *9*, 2006.

Kosar, Kevin R. (2005). *Failing grades : the federal politics of education standards / Kevin R. Kosar*: Boulder, Colo. : Lynne Rienner Publishers, 2005.

McDonnell, Lorraine M. (2005). No Child Left Behind and the Federal Role in Education: Evolution or Revolution? *Peabody Journal of Education (0161956X), 80*(2), 19-38. doi: 10.1207/S15327930pje8002_2

McKay, Robert B. (1958). An American Constitutional Law Reader; A Collection of Materials on American Constitutional Law, including Excerpts from Supreme Court Opinions, Commentary by Students of the Constitutional Process, and Supplemental Observations of the Editor 204 Constitutional Limitations on the Power of Government

Medina, Noe J., & Neill, D. Monty. (1990). *Fallout from the testing explosion : how 100 million standardized exams undermine equity and excellence in America's public schools / by*

Noe Medina and D. Monty Neill ; with the staff of the National Center for Fair and Open Testing (FairTest): Cambridge, MA : National Center for Fair & Open Testing (FairTest), 1990.Rev. 3rd ed.

Mueller, Jennifer. (2001). Facing the Unhappy Day: Three Aspects of the High Stakes Testing Movement. *Kansas Journal of Law & Public Policy, 11*, 201.

National Commission on Excellence in, Education. (1983). A nation at risk: the imperative for educational reform.

O'Connell v. Superior Court, No. No. A113933, 47 147 (Cal: Court of Appeal, 1st Appellate Dist., 4th Div. 2006).

O'Neill, Paul T. (2003). High Stakes Testing Law and Litigation (pp. 623).

Palardy, J. Michael. (1984). *Education, 104*(4), 401.

Popham, W. James. (2001). *The truth about testing : an educator's call to action / W. James Popham*: Alexandria, Va. : Association for Supervision and Curriculum Development, c2001.

Position Statement of the American Educational Research Association concerning High-Stakes Testing in PreK-12 Education. (2000). *Educational Researcher, 29*(8), 24-25. doi: 10.2307/1176634

Quigley, William P. (2000). Due Process Rights of Grade School Students Subjected to High-Stakes Testing. *B.U. Pub. Int. L.J., 10*, 284-319.

Quigley, William P. (2001). Due process rights of grade school students subjected to high-stakes testing. *Boston University Public Interest Law Journal, 10*(2), 284-319.

Rankins v. Louisiana State Bd. of Elementary & Secondary Educ., 637 548 (Court of Appeals, 1st Circuit 1994).

Ravitch, D. (2013). *Reign of Error: The Hoax of the Privatization Movement and the Danger to America's Public Schools*. Random House Digital, Inc.

Recent Court Decisions and Legislation Impacting Juveniles Recent Court Decisions. (2006). (Vol. 10).

Robelen, Erik W. (2001). Most Conservatives Are Backing the President--for Now. *Education Week, 20*(21), 21.

Saleh, Matthew. (2011). Modernizing "San Antonio Independent School District v. Rodriguez": How Evolving Supreme Court Jurisprudence Changes the Face of Education Finance Litigation. *Journal of Education Finance, 37*(2), 99-129.

San Antonio Independent School Dist. v. Rodriguez, No. No. 71-1332, 411 1 (Supreme Court 1973).

Scheelhaase v. Woodbury Central Community Sch. Dist, No. Civ. No. 71-C-3029-W, 349 988 (Dist. Court, ND Iowa 1972).

Scheelhaase v. Woodbury Central Community Sch. Dist, No. No. 73-1067, 488 237 (Court of Appeals, 8th Circuit 1973).

Serow, Robert C, & Davies, James J. (1982). Resources and outcomes of minimum competency testing as measures of equality of educational opportunity. *American Educational Research Journal, 19*(4), 529-539.

Stratman, Dave. (2001). A Call For "Mass Refusal". *Social Policy, 32*(2), 55-60.

Test Opposition Surges Across the Nation. (2013). Retrieved June 7, 2013, from http://www.fairtest.org/test-opposition-surges-across-nation

The use of tests when making high-stakes decisions for students : a resource guide for educators and policy-makers / U.S. Department of Education, Office for Civil Rights. (2000). [Washington, D.C.] : The Office, [2000].

U.S. Const. Amend. XIV, Sect. 1

U.S. Const. Amend. XIV, Sect. 2

U.S. Const. Art. VI

Vinovskis, Maris. (2009). *From A nation at risk to No Child Left Behind : national education goals and the creation of federal education policy*. Teachers College Press: New York.

Welner, Kevin. (2010). Education Rights and Classroom-Based Litigation: Shifting the Boundaries of Evidence. *Review of Research in Education, 34*(1), 85-112.

What's Wrong With Standardized Tests? (2012). Retrieved September 15, 2013, from http://www.fairtest.org/facts/whatwron.htm

Williams, S. M., & Houck, E. A. (2013). The Life and Death of Desegregation Policy in Wake County Public School System and Charlotte-Mecklenburg Schools. *Education and Urban Society*.

Wilson, P. E. (1964). Brown v. Board of Education revisited. *University of Kansas Law Review., 12*, 507-507.