

## ABSTRACT

### No Child Left Behind Act: The Impact of Standards-based Accountability

Eric Shin

Director: Nicholas Benson, Ph.D.

What was the impact of the Standards-Based Accountability system of the No Child Left Behind Act? SBA required states to produce a set of standards for standardized testing, and if these standards were not met, sanctions were imposed on schools. This was enforced through the Adequate Yearly Progress Report which was created under the NCLB. This thesis explores the impact of SBA implementation on the educational quality, notably teaching to the test and curriculum narrowing. Furthermore, SBA led to increased rates of test anxiety. This paper concludes by proposing recommendations to the accountability system and exploring the changes brought forth under the Every Student Succeeds Act.

APPROVED BY DIRECTOR OF HONORS THESIS:

---

Dr. Nicholas Benson, Department of Education

APPROVED BY THE HONORS PROGRAM:

---

Dr. Elizabeth Corey, Director

DATE: \_\_\_\_\_

NO CHILD LEFT BEHIND ACT: THE IMPACT OF STANDARDS-BASED  
ACCOUNTABILITY

A Thesis Submitted to the Faculty of  
Baylor University  
In Partial Fulfillment of the Requirements for the  
Honors Program

By  
Eric Shin

Waco, Texas

May 2022

## TABLE OF CONTENTS

List of Tables.....	iii
List of Figures.....	iv
Chapter One: Introduction.....	1
Chapter Two: The History of NCLB and Standards-based Accountability.....	3
Chapter Three: Teaching to the Test and Curriculum Narrowing.....	17
Chapter Four: Increased Stress and Fear.....	38
Chapter Five: Evaluating SBA and Recommendations for Improvement.....	50
Chapter Six: Conclusion.....	62
References.....	64

## LIST OF TABLES

Table 1, Percentage of Schools Identified for Improvement under AYP.....	14
Table 2, AYP Standards in TX, NY, and MA.....	24
Table 3, Influence of High-Stakes Testing on Curricular Instruction.....	28
Table 4, Instructional Hours and School Improvement Status.....	32

## LIST OF FIGURES

Figure 1, Sanctions for <i>Missing AYP</i> .....	12
Figure 2, Model of Teaching to the Test.....	18
Figure 3, Test Anxiety Model.....	39
Figure 4, Test Anxiety between Classroom and NCLB-mandated Tests.....	42

## CHAPTER ONE

### Introduction

By the beginning of the 21<sup>st</sup> century, there was a growing public belief that the United States' education was falling behind the rest of the world. Hoping to energize the education system, George Bush passed the No Child Left Behind (NCLB) of 2001 with overwhelming bipartisan support. The act introduced Standards-Based Accountability (SBA), an overarching punitive system that sought to identify and sanction underperforming schools.

To fully understand NCLB, the history leading up to its passing is examined. In the 20<sup>th</sup> century, there was a general trend of federalization in standardized testing. From the Elementary and Secondary Education Act of 1965 to the Improving America's Schools Act of 1994, standardized testing and the accountability system has become increasingly controlled by the federal government, a trend that continued into NCLB.

The impact of the Standards-Based Accountability System is then examined. This was unprecedented in its scope compared to previous acts. From the firing of the staff to school restructuring, the consequences of not meeting state standards were significant. Under pressure from both the public and administrators, rates of teaching to the test increased. Items most likely to appear on tests became the focus of the curriculum, preventing students from receiving a more comprehensive education. This was facilitated by misalignment between state standards and the assessments, meaning certain test items were overrepresented and thus predictable. The accountability system's focus on math and

English Language Arts (ELA) also led to the erosion in teaching of other subjects in a process known as curriculum narrowing.

Another consequence of SBA's harsh sanctions was the pressure it placed on students. Rates of test anxiety increased under NCLB, and some students experienced anxiety to a level so debilitating that it could significantly harm their performance on the exam. The high stakes of the exams also increased anxiety in teachers, who worried about the consequences if their students did not meet standards. Several intervention methods, for example the *TestEdge* system, can help in reducing test anxiety.

The final chapter explains the benefits of standardized testing and why it should not be replaced by alternatives such as Portfolio Assessments. Some benefits include ease of use, its ability to compare students and teachers, and the standardization that it can provide. Some ways to improve NCLB's accountability system are then proposed. By making sanctions less punitive, widening the subjects considered under SBA, and improving alignment between standards and assessments, the consequences of the accountability system can be mitigated. Some of these changes were implemented into the Every Student Succeeds Act of 2014 which will help to reduce issues like teaching to the test and curriculum narrowing. In all, this thesis explains how the No Child Left Behind Act fits into the history of standardized testing in the U.S., the consequences of its unprecedented standards-based accountability system, and how the Every Student Succeeds Act introduces reforms that to resolve NCLB's consequences.



## CHAPTER TWO

### The History Behind NCLB and Standards-Based Accountability

The No Child Left Behind (NCLB) Act of 2001 was a landmark law passed by President George W. Bush to close the achievement gap and raise America's educational standing in the world. The act was novel in the extent of its standards-based accountability (SBA) system. However, to truly understand the significance of NCLB and its accountability system, the history leading up to its passing must first be evaluated. Jesse H. Rhodes in *An Education in Politics: The Origins and Evolution of No Child Left Behind* states that in stark contrast to the NCLB's focus on federal standards and accountability, the vast history of American education was a "very decentralized affair", with states and localities being responsible for managing education (26-27). This was due to the widespread belief that education was primarily a state's right. This belief continued into the 1940's with considerable public resistance against federalization. The resistance can be summarized with the three "R's": race, religion, and reds. Civil rights activists were "reluctant to endorse federal education bills that benefitted segregated school districts" (29) and anti-Catholics opposed giving aid to Catholic schools. Federal aid could not be selectively given to schools on factors such as race and religion, leading many to oppose federalization of the education system. Also important was the Cold War and the resulting public fear of communism. The public was concerned of government control of education as it was considered to be "communist", further opposing government involvement (29).

These factors meant that for much of American history, the federal government did not have much discretion over the education system.

The federal government's hands-off attitude towards the education system did not last. Doubts over the quality of American education first arose during the Cold War. According to David A. Gamson, Kathryn McDermott, and Douglas Reed in "The Elementary and Secondary Education at Fifty: Aspirations, Effects, and Limitations," Russia's progress in space sparked a "Crisis in Education" with *Life* magazine describing the American education standards as "appalling" and "shockingly low" (5). Comparisons were even drawn between hard-working Soviet students and their "lazy" American counterparts. Due to the public's fears of inferiority to the Soviets, the federal government was able to increase its influence on the education system. In 1958, Congress passed the National Defense Education Act (NDEA) which gave federal aid to states for education, giving loans to students and funds to improve science, math, and language education (US Dept. of Education). Ultimately, though, the NDEA did not significantly expand the federal government's influence over the education system as federal aid constituted "less than 2 percent of all education spending" (Rhodes 30).

Under the Presidency of Lyndon B. Johnson, a widespread shift occurred in favor of federal involvement in the education system. As David Casalaspi explains in "The Making of a 'Legislative Miracle': The Elementary and Secondary Education Act of 1965," the passage of the Elementary and Secondary Education Act (ESEA) in 1965 was motivated by the desire to close the socioeconomic and racial achievement gap, an integral aspect of Johnson's War on Poverty (254). President Johnson was the son of teachers and had taught in impoverished school districts and realized that education was critical for the

American Dream. The preceding President Kennedy, perhaps due to his privileged upbringing, did not share Johnson's conviction on this issue (253). Also relevant was the racial landscape. Only a decade prior, *Brown v. Board of Education* brought racial inequality in the education system to the forefront of American consciousness, and ESEA in part was a response to these calls for equality. (Rhodes 30). The ESEA consisted of five titles, the most important being Title I. Title I allocated funds to disadvantaged students to close the achievement gap, and with a budget of \$1 billion, it signaled a crucial step in federal involvement of education (Gamson et al. 9). Title I was the first time the federal government allocated a significant amount of money towards education. Its significance cannot be overstated as even today, Title I funding is used for lower income students.

Although the ESEA worked to expand federal jurisdiction in education, it still did not allow for the vast oversight of today's federal government. Due to the history of a decentralized education system, the ESEA "delegated substantial discretion to states and localities to design and implement compensatory education programs" (Rhodes 33) which crucially meant that there was no accountability system in place to check local governments' use of federal funds. This would lead to variation in schools' rigor and administrative practices. Without accountability, what would stop some schools in misusing Title I funds? Indeed, some school districts inappropriately spent money on "band uniforms" and "swimming pools" and many districts did not prioritize educational needs of poor students. Not surprisingly, the massive Title I spending under ESEA did not significantly raise performance of children in poverty (Gamson 15). ESEA also promoted "pull-out programs" in which disadvantaged students were given separate instruction from other students, meaning that Title I funds only meaningfully impacted a subset of a receiving school's

student population and worked to fragment the recipients of aid (Rhodes 33). This called into question the fairness of the act as non-disadvantaged students did not experience the benefits of the Title I funds.

The bombshell report “A Nation at Risk: The Imperative for Educational Reform” by the National Commission on Excellence in Education in 1983 conveyed to the public that the American education system was falling behind, bluntly stating that “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” (112). Evidently, ESEA had not alleviated the fears of inferiority first brought on by Sputnik. The report stated that in a world in which the Japanese make better cars, South Koreans make better steel mills, and Germans make better machine tools, America is losing its competitive edge and must dedicate [itself] to the reform of [its] education system for the benefit of all” (114). The report lists a plethora of findings on why the nation is at risk, a few examples being a large functionally illiterate adult population, a steady decline in science achievement scores, and lower standardized test scores than when Sputnik was launched (116). One recommendation the report makes is the implementation of a more rigorous standards and higher expectations for academic performance, something the ESEA critically lacked (125). John W. Hunt argues in “A Nation at Risk and No Child Left behind Déjà Vu for Administrators?” that A Nation at Risk had a widespread impact on the education policies of the decades following its publication. It has sparked a “standards movement” that has shifted public focus from the activities of teachers to the achievement of students (583). This prioritization in student performance is crucial for the push for an accountability system as implemented in the Improving America’s Schools Act (IASA).

The Improving America's Schools Act of 1994 was a reauthorization of the ESEA that included critical revisions, notably in the accountability system. According to H.R.6 of the 103<sup>rd</sup> Congress, the IASA "requires any State desiring to receive a grant... to submit State plans that are coordinated with other programs under ESEA" and "requires each State plan to describe challenging standards for all children", setting challenging academic and student performance standards in math and English (H.R.6 1994). In contrast to the ESEA which delegated use of federal spending to the states, the IASA required states to set challenging educational standards to receive the Title I grant in the first place. Furthermore, if schools could not meet standards, they were required to devise a school improvement plan so they could meet performance standards (Gamson et al. 16). Michael Puma and Darrell Drury assert in *Exploring New Directions Title I in the Year 2000* that along with standards-based reform, IASA dropped the poverty rate threshold for operating schoolwide programs from 75% to 50%, increasing the number of schools eligible, and increased flexibility by allowing states to waive "certain federal requirements if they interfere with school improvement" (4-5). From the beginning of American education to the IASA, it is evident that the overall trend is increased federal power over the education system and an expanding accountability system.

The IASA was passed in an era of broad education reform. Student achievement scores had stagnated by the early 1990's, and large-scale surveys found that disadvantaged students were given less challenging materials leading to their dissatisfaction (Rhodes 99). This featured prominently in the 1992 Presidential election in which both Bush and Clinton advocated for widespread education reform. After Clinton's election, he would pass Goals 2000: Educate America Act, which outlined 8 national goals to be achieved by the year

2000. “20 USC 5812: National Education Goals” states that the goals include raising graduation rates to 90% and student mastery of subjects such as mathematics and English. According to Richard W. Riley in “The Improving America’s Schools Act and Elementary and Secondary Education Reform,” the Goals 2000 Act and IASA worked in conjunction, with the IASA providing the goals and framework necessary to implement Goals 2000 (520). This accountability system under IASA was therefore expected to decrease inequality in education and raise America’s standing in the world.

However, the IASA could not implement a meaningful standards-based accountability system because it gave too much control to the state and local governments. By 1998, the IASA required that states set content standards specifying what students should know and performance standards specifying achievement, yet less than half had done so (Rhodes 133). Also, because of “uneven collection and reporting of student performance data and inconsistent use of exemptions to exclude certain students from accountability tests”, some states would overestimate the number of struggling schools, overwhelming the federal government who had to provide funding, while others would underestimate struggling schools (Rhodes 133). In essence, leaving states to implement an accountability system resulted in failure and the IASA’s goal of implementing high standards for students was not met.

As the governor of Texas in the 1990’s, George W. Bush saw an apparent improvement in student performance. According to Walt Haney in “The Myth of the Texas Miracle in Education,” the roots of this miracle lay in the introduction of Texas Assessment of Academic Skills (TAAS), a standardized test of mathematics, reading, and writing. TAAS was administered to students from grades 3 to 10 with the exception of grade 9

Furthermore, TAAS showcased a new accountability system that punished schools for not meeting standards, with a school's "reputation, funding, and continued existence [depending] on students' performance on TAAS" (Haney 12). From 1994 to 1998, TAAS and the accountability system seemed to show massive improvement in student performance, with TAAS passing rates rising from 57% to 77% in that time period (14). Dubbed the "Texas Miracle", TAAS and the accountability system's apparent success was a foreshadowing of the coming No Child Left Behind Act. Although the veracity of the Texas Miracle has been questioned in following years, TAAS was nonetheless a significant factor in President George W. Bush wanting to expand the power of the federal government in education.

The disappointment of the IASA to implement meaningful reform and the success of the "Texas Miracle" led to the creation of the No Child Left Behind Act of 2001, a reauthorization of the earlier Elementary and Secondary Education Act. In *No Child Left Behind and the Reduction of the Achievement Gap: Sociological Perspectives on Federal Educational Policy*, Alan R. Sadovnik et al. explain that to fulfill the goals of the earlier Goals 2000 and close the achievement gap, President Bush pushed for annual testing for accountability. NCLB differed from earlier acts by tying a school district's performance as measured by standardized test to federal funds; in other words, schools not meeting standards would not receive money (Sadovnik et al. 16). States could no longer blow off the implementation of an accountability system as the loss of funding from not meeting standards could be devastating. As Wang et al. explain in "Controversies of Standardized Assessment in School Accountability Reform: A Critical Synthesis of Multidisciplinary Research Evidence", the NCLB's main goal is to close the achievement gap between

students regardless of disabilities, socioeconomic status, or race, and mainly accomplishes this using high-stakes consequences, a component of the accountability system (306). With such lofty goals, the NCLB received both political and public support for SBA and standardized testing.

In order to improve achievement for all students, NCLB separates the student population into four subgroups: economically disadvantaged, racial minorities, students with disabilities, and students with limited English proficiency (Dillon and Rotherham 2). The use of subgroups allowed states to determine if some subgroups were underachieving compared to others, something that could not be easily elucidated prior to NCLB. During the 2005-2006 school year, Connecticut met the state reading assessment proficiency requirement, but a close analysis showed that disabled students and English-deficient students scored below the requirements (Dillon and Rotherham 2).

A close reading of the No Child Left Behind Act reveals the extent of the changes.

Section I-A of NCLB establishes requirements for:

(1) yearly testing and assessments of student performance; (2) State standards for and assessments of Adequate Yearly Progress (AYP); (3) local educational agency (LEA) identification of schools for improvement and corrective actions; (4) reporting to parents and the public on school performance and teacher quality; (5) eligibility requirements for schoolwide programs; and (6) increased qualifications of teachers and paraprofessionals. (H.R. 1 107<sup>th</sup> Congress)

The establishment of yearly standardized tests and new standards under AYP meant that NCLB established a far-reaching system of accountability, a system that eclipses ESEA and ISIA. Laura S. Hamilton et al. explain in *Standards-Based Accountability Under No Child Left Behind: Experiences of Teachers and Administrators in Three States* that the term Standards-Based Accountability (SBA) is a combination of three ideas: academic standards, standardized assessments, and accountability for student outcomes (2).



Although these three ideas were present in the American education system, for example the IASA setting academic standards, the NCLB's integration of the three ideas into one cohesive system was novel (2).

SBA established a "multilevel, multistep feedback mechanism" with the goal of informing school districts if students were taught effectively. Performance standards dictated what students should know at a given age, and school districts made their curriculum based on these standards. The standardized assessments were created with these standards in mind, and performance of students on these assessments indicated school success or failure, thus establishing the "feedback mechanism" (Hamilton et al. 3). Therefore, SBA would ideally give concrete goals that schools could work towards, and the standardized tests served as an objective assessment to evaluate if they have met these goals. However, as will be explained in following chapters, NCLB's ambitious goal of narrowing the achievement gap and raising student achievement was not as successful as intended.

Under SBA, states established Adequate Yearly Progress (AYP) which measured the school's performance compared to the standardized tests. For a state to make AYP, it must: reach performance standards, meet participation requirements, and meet performance requirements for "an additional indicator of academic performance determined by the state" (Dillon and Rotherham 2). Although not part of the original law, states could use an "index score" as a performance standard that recognized schools for improving performance, even if the "proficiency" threshold is not met. Furthermore, NCLB required a 95% participation requirement to prevent schools from not testing low-scoring students to inflate scores. If a school were to not meet the 95% requirement, they

would miss AYP. The “additional indicator of performance” for most schools was the graduation rate (2-3).

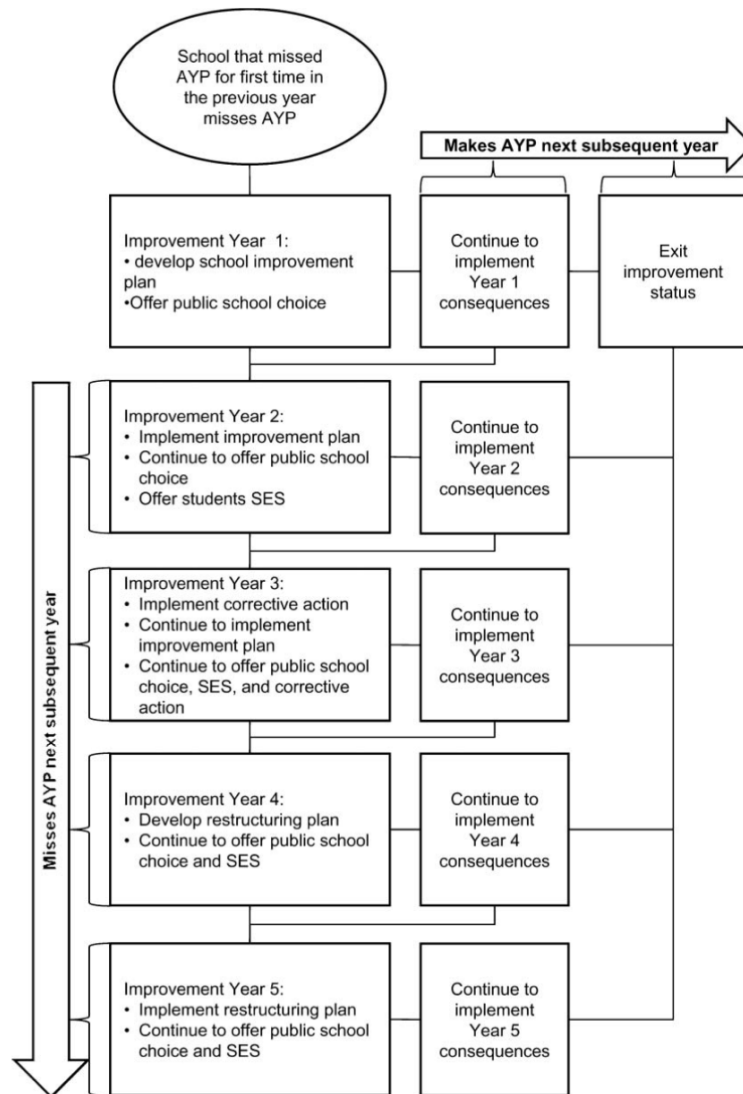


Fig. 1. Sanctions for Missing AYP (Forte 81).

NCLB’s standards-based accountability notably differed from previous acts by establishing a “punishment system” for schools that fail to make AYP, summarized in Figure 1 by Forte. These sanctions could be severe, discouraging schools from half-

heartedly implementing education changes. The Wisconsin Department of Public Education outlines these sanctions for Schools Identified for Improvement (SIFI) status. There were no sanctions in the first year a school misses AYP, but after a school missed AYP for two consecutive years, the district had to make a 2-year improvement plan and allocate 20% of Title I funds to fund transportation of students electing to transfer. After 3 years of missing AYP, the district had to “make supplemental education available and notify parents” as well as continuing to allocate Title I funding for transfer student transportation. Missing 4 years of AYP triggered a “Corrective Action Status” in which districts implemented harsher measures such as replacing staff responsible for failure or instituting a new curriculum. After 5 years of missing AYP, the school was placed in “Restructuring Status” and implemented drastic reform such as making the school a charter school or replacing most of the staff. Finally, 6 years of missing AYP required schools to implement the reform planned in the “Restructuring Status” phase (Wisconsin Department of Public Education 1-2). As President Bush explained in his speech at General Phillip Kearny School in 2009,

The philosophy behind the law is pretty straightforward: Local schools remain under local control. In exchange for federal dollars, however, we expect results. We're spending money on schools, and shouldn't we determine whether or not the money we're spending is yielding the results society expects? So states set standards. One reason this school makes sense is because you have a principal who sets high standards, keeps that bar high. And we hold schools accountable for meeting the standards. There -- we set an historic goal, and that is to -- every child should learn to read and do math at grade level by 2014. (Bush speech 2009)

In other words, these stringent sanctions were partly to justify the massive federal spending under Title I of the NCLB. The ambitious goal of 100% math and reading proficiency by 2014 in return required ambitious sanctions not seen previously in America's education system.

Along with sanctions, NCLB set up a awards system for schools that met expectations under AYP. State governments were responsible for enacting this system which results in variation among states. California selected around 5% of public schools for a monetary award for meeting AYP. Georgia, having already implemented a pay-for-performance program before NCLB, expanded this program by awarding schools based on making AYP and the percentile of its students meeting standards, along with a monetary reward for schools making AYP for three consecutive years. Pennsylvania also implemented a rewards program in 2004 (Hamilton et al. 32-33). The limited number of schools receiving awards, along with a relative lack of standardization, makes it clear that No Child Left Behind emphasized punishment over rewards. In other words, the potential for funding loss and reconstruction motivated schools to improve student performance, not financial rewards that few schools received.

	Population	State-Identified		Research-Identified		Non-Identified	
	2009 and 2010	2009	2010	2009	2010	2009	2010
Number of schools	1,059	105	75	427	422	527	562
Percentage of elementary	59.6	20.0	30.7	65.6	63.7	62.6	60.3
Percentage middle	21.4	49.5	28.0	16.9	19.4	19.5	22.1
Percentage high	19.0	30.5	41.3	17.6	16.8	17.8	17.6

Table 1. Percentage of schools identified for improvement under AYP (Hochbein 278).

This begs the question: was Adequate yearly Progress implemented fairly and effectively? In “Gamed by the System: Adequate Yearly Progress as an Indicator of Persistently Low-Achieving School Performance”, Craig Hochbein et al. analyzed data from all public elementary, middle, and high schools in a midwestern state and gathered the school’s measured annual objectives (AMOs) and AYP status. They defined “state-

identified groups” as schools missing AYP for 3 consecutive years (276). Although the majority of schools studied (60%) were elementary schools, they only constituted 20% of state-identified schools in 2009. In contrast, high schools were only 19% of the sample but constituted 30% of state-identified schools (278). They also noticed that the number of AMO targets were reduced in Elementary schools. Putting these two data sets together, Hochbein et al. concluded that AYP systematically under-identified Elementary Schools as low-performing and likewise overidentified High Schools. This error resulted because elementary schools’ reduced number of AMO’s, likely due to their smaller population size, resulted in easier AYP requirements compared to high schools (284). The significant bias in AYP identification had major impact in the state studied. In 2009-2010, 21 schools, 18 of which were high schools, were subject to significant reform where principals and a significant number of staff were replaced. No elementary schools underwent such reform (281).

Further research has confirmed Hochbein’s claim of biased identification by AYP. Daniel Betebenner et al. in “The Impact of Performance Level Misclassification on the Accuracy and Precision of Percent at Performance Level Measures” state that literature that downplays bias in AYP are “plagued by inconsistencies in both notation and terminology” (120). They were able to quantify bias using techniques more accurately from randomized response designs, and by using 2005 state assessment data, they concluded that low-achieving schools overestimated the number of proficient students; more students were labeled “proficient” than they should be (130). On the other hand, high-performing schools underestimated the number of proficient students. The authors further conclude that the stringency of a state’s proficiency can “differentially impact high- and low-

performing schools due to differential bias resulting from misclassification” (Betebenner et al. 132). The results of this study cast doubt on the validity of AYP and the accountability system and suggests it may not be the objective standard it claims to be.

The No Child Left Behind Act, with its goals of improving student performance and decreasing the achievement gap, implemented a standards-based accountability system that touted equality and standardization. It is thus expected and crucial that Adequate Yearly Progress can identify underperforming schools accurately without bias. After all, why have a standards-based system that isn’t standard? Both Hochbein et al’s and Shang et al’s research cast doubt on its fairness, instead arguing that AYP is plagued with bias, from overrepresentation of struggling high schools to misclassification of the number of proficient students. It becomes evident that one cannot take NCLB’s touted benefits at face value, and only through detailed analysis and literature review can a fair judgement on its efficacy and fairness be truly assessed.

## CHAPTER THREE

### Teaching to the Test and Curriculum Narrowing

A reform as big as the No Child Left Behind Act inevitably draws its share of criticism. One criticism shared by educators and researchers alike is that NCLB promoted teaching to the test. Before we proceed, teaching to the test must first be defined. According to “Teaching to the Test?” by W. James Popham, teaching to the test occurs when teachers provide actual test questions as part of the practice material or provide “clone items” that are difficult to distinguish from actual test items (16). The goal of the curriculum then becomes scoring well on the standardized exams, not providing a balanced education to students. The quality of the education system is thus sacrificed when teaching to the test occurs. A standardized exam cannot realistically test all the state standards; therefore, it must test only a sample which are oftentimes not representative. Popham gives an example of a vocabulary quiz that contains 20 items out of a larger collection of 200 words. If a teacher were to tailor their teaching specifically to those 20 words, then this would be *item teaching*, also known as teaching to the test. This seeks to artificially raise the scores of students without an increase in their proficiency (16). After all, when a student performs well on a test, it is inferred that he or she has mastered the overall material, which is nullified when specific test items are taught to the test.

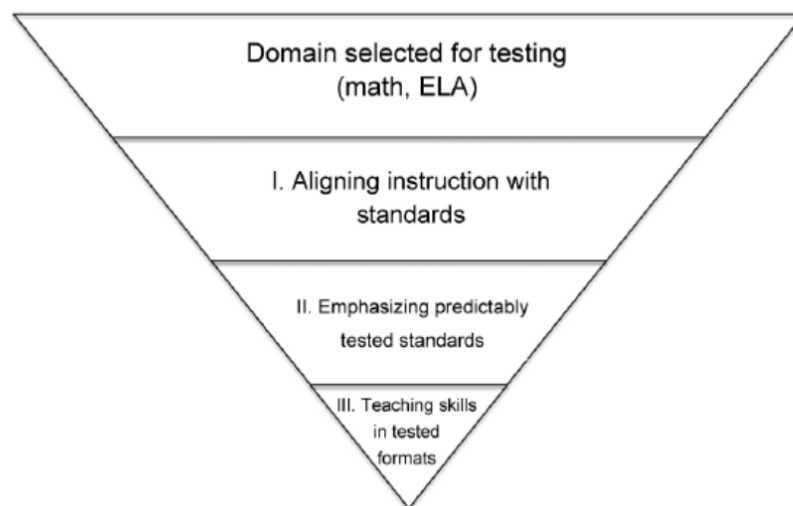


Fig. 2. Model of Teaching to the Test (Jennings and Bearak 382).

Teaching to the test is not black-or-white, instead existing on a continuum. In “‘Teaching to the Test’ in the NCLB Era: How Test Predictability Affects Our Understanding of Student Performance,” Jennifer L. Jennings and Johnathan Marc Bearak state that teaching to the test progressively narrows from the domain of interest, as shown in Figure 2. In stage I, subjects are relocated to align with state standards. Furthermore, when teachers add subjects to the curriculum because it is likely to be on the standardized exam, it is usually followed by the omission of other materials (382). Students are therefore not necessarily learning more. There also exists a mismatch in *alignment* between state assessments and state standards; this issue is increasingly important in the NCLB discussion and will be elaborated further in the coming paragraphs. In stage II of teaching to the test, teachers emphasize certain state standards that are overrepresented on state tests. Due to misalignment between standards and tests, “state test results can become inflated when a small fraction of state standards is predictably tested over time” (Jennings and Bearak 383) which leads to educators emphasizing high-yield standards and thus reducing



the validity of the tested students' proficiency of the overall state standards. Finally, in stage III, educators teach items in the same format as they appear in state exams. Test items often vary very little year-to-year so educators can teach specifically to that format, thus bringing into question if students are truly learning the material.

Teaching to the test was driven by the No Child Left Behind Act. As explained in the previous chapter, NCLB mandated yearly testing on ELA and mathematics for students in grades 3 through 8, and a certain percentage in each subgroup (gender, race, etc.) had to be deemed "proficient" under Adequate Yearly Progress. If a school failed to meet AYP, they were met with harsh sanctions that can result in loss of funding and school restructuring. Daniel Koretz in "Adapting Educational Measurement to the Demands of Test-Based Accountability" explains that these sanctions placed substantial pressure on school districts, and educators often resorted to emphasizing tested materials in place of teaching a broader curriculum, also known as teaching to the test. Previous acts such as the IASA of 1994 had an accountability system, but the NCLB was novel in the extent of its accountability and sanctions. Abigail B. Brown and Jack W. Clift in "The Unequal Effect of Adequate Yearly Progress: Evidence From School Visits" find that some schools encouraged "a shift away from open-ended work that is designed to encourage higher level thinking and toward a much greater emphasis on the use of multiple-choice questions in the classroom..." (783). One teacher interviewed stated that if they did not teach to the test, their students scored poorly. Crucially, she clarified that what she was pressured to do was teach to the test itself, not to the standards (Brown and Clift 783). Teaching to the standards is not inherently bad. After all, standards tell teachers what the students should be proficient in. However, when teachers prioritize overrepresented standards and test formats, the main

goal of the curriculum shifts from providing a balanced education to scoring well on the tests.

Although teaching to the test surged under NCLB, research shows it was present years before the act's passing. In *The Validity of Gains in Scores on the Kentucky Instructional Results Information System (KIRIS)*, Daniel Koretz explores the large increase in test scores shown under KIRIS, a 1990's reform program in Kentucky that implemented SBA and sanctions before NCLB. From 1992 to 1997, all grades had significant improvements in student categories for all subjects, leading to KIRIS being considered a success in improving student proficiency (23). However, further research reveals that KIRIS gains were "implausibly large" (113) and were significantly inflated compared to external assessments. Crucially, students did better on reused items compared to new items, suggesting that item-specific teaching to the test occurred (115). Surveyed teachers also believed the increase in test scores was due to test familiarization and narrowing the curriculum to high-yield subjects, not through gains in knowledge (110). Just as KIRIS promoted teaching to the test in Kentucky through standards-based accountability, the No Child Left Behind Act also promoted teaching to the test on a nationwide scale.

The No Child Left Behind Act mandated end-of-the-year assessments to determine student performance. To prepare students for these assessments and gauge their progress, teachers used *interim* or *benchmark assessments*. In "What the Marketplace Has Brought Us: Item-by-Item Teaching with Little Instructional Insight," Lorrie A. Shepard defines interim assessments as "test products designed to be administered every quarter or every 6 weeks" that supplement NCLB's high-stakes exams. Ideally, interim assignments would

find gaps in student knowledge so teachers could adjust their curricula dynamically (Shepard 246). In reality, however, they focused most on “bubble kids”, defined as those who are immediately below the proficiency cutoff, rather than the entire student population. In addition, interim assignments were used to prepare for the high-stakes exams rather than identify gaps in student knowledge. Shepard states that “interim assessments results appeared to be used, item by item, to reteach the steps in the problems that were missed” with “no mention of attending to the underlying concept” (255). Clearly, teachers used these exams to teach test items, not concepts. By definition, this is teaching to the test and conflicts with interim assessments’ goal of filling in students’ knowledge gaps. Shepard’s conclusions are significant because it highlights the pervasiveness of teaching to the test.

One underlying factor in teaching to the test is *alignment*. According to Betsy J. Case et al., alignment is the “degree to which the components of an education system—such as standards, curricula, assessments, and instruction—work together to achieve desired goals” (2). A particular area of focus in the case of NCLB is the alignment between state standards and the standardized assessment. It asks, “does the assessment assess what it is *supposed* to”? For example, if a standard is that 3<sup>rd</sup> graders should know how to do addition, does the test actually test on addition and not, say, subtraction? Case et al. define this kind of alignment as “horizontal alignment” (9). Alignment is especially relevant in NCLB because the standards-based accountability system clearly establishes standards for student achievement. There are 3 general ways to create alignment: sequential development, expert review, and document analyses (5). Regardless of the alignment system used, it is important that any test, not just those mandated under NCLB, features alignment between the test and the standards, because misalignment would harm the validity of the test.

Alignment also occurs between the curriculum and the assessment. It asks, “does the curriculum cover what is actually on the test”? By setting clearly defined standards, NCLB sought to promote this curriculum-assessment alignment and thus improve student performance. Because teachers are motivated to align their curriculum in order to avoid sanctions under AYP, alignment should in theory increase over time. This is not always the case. Morgan S. Polikoff in “Instructional Alignment under No Child Left Behind” uses self-reported data on teacher instruction and compares this with the content of standards and assessments, allowing her to determine the extent of the alignment between the curriculum and the standards and assessments in English, math, and science (342). Results of the study show that instructional alignment differed significantly between subjects, with large increases of 0.65 SD in mathematics, a smaller 0.1-0.2 SD increase in English, and decreases in alignment in science. Polikoff attributes this to the more concrete and well-defined standards for mathematics compared to English, and alignment in science likely decreased as it is a less significant aspect of the accountability system. Also notable is that the overall increases in alignment were small compared to the “near unanimity with which teachers report efforts to improve their alignment across subjects and grades in surveys” (361). It is evident that even with effort from teachers, the level of alignment did not increase meaningfully.

Misalignment between assessments and standards predates the No Child Left Behind Act. In *Redesigning Accountability Systems for Education*, Fuhrman and Elmore explore the relationship between standards and assessments. Ideally, standards “provide guidance to textbook and materials publishers, test developers, and professional development providers” that ensures alignment between standards and classroom

instruction (Fuhrman and Elmore 97). With proper alignment, educators know whether students have achieved expectations. For proper alignment, tests should match the standards in both *depth* (extent of abilities tested) and *breadth* (range of skills). The Third International Mathematics and Science Study (TIMSS) of 1997, an extensive cross-national study, and Achieve's Benchmarking Initiative of 1998 were two methodologies used to design an "alignment protocol". This protocol analyzed four aspects to measure alignment: 1. *Content centrality* (the alignment between the test and standards), 2. *Performance centrality* (the alignment of cognitive demand, e.g., "compare" or "analyze", between the test and standards) 3. *Challenge* (whether doing well requires mastering challenging material), and 4. *Balance and range* (whether the test is balanced compared to standards and covers the breadth of the standards). An ideal testing system should meet all four of these criteria.

The TIMSS and Achieve studies revealed that "tests do not measure what the standards expect, either in the content and performances expected of students or in the breadth of knowledge and skills outlined in the standards" (Fuhrman 111). Furthermore, tests were predictable in that the least cognitively complex expectations were disproportionally represented. If assessments are not aligned with standards, how valid are these exams? A student could perform well on a test without mastering the state-mandated standards. This misalignment and predictability of standardized assessments sets up the basis for teaching to the test because educators have an incentive to deviate from standards to raise student scores. As Fuhrman and Elmore explain, "some will likely focus only on the material on the test, thereby shortchanging their students of higher-level knowledge and skills..." (112). The introduction of standards in standardized testing has the benefit of

providing guidance to educators and test writers, but in return, misalignment with assessments is a significant area of concern.

Description of Standard Coverage in New York (NY), Texas (TX), and Massachusetts (MA) by Subject and Grade, 2006 to 2009						
Variable	ELA			Math		
	NY	TX	MA	NY	TX	MA
Number of standards	59	36	36	41	19	18
Percentage of standards tested in 2009	41	94	79	27	79	49
Percentage of standards that ever appeared on state tests (2006–2009)	60	100	97	41	93	62
Percentage of standards students must master to earn 50% of test points	15	32	18	13	22	19

*Note.* Table reports means across all grades. ELA = English language arts.

Table 2. AYP Standards in TX, NY, and MA (Jennings and Bearak 384).

Evidence points to the continuing of this misalignment into the NCLB era. Jennings and Bearak examined ELA and math scores from grades 3 to 8 in New York, Texas, and Massachusetts to determine how closely state tests matched with state standards. This data was *item-level*, meaning each question was analyzed against the state standards and student performance was determined (383). As evident in Table 2 above, from 2006–2009, Texas tested 100% of its standards for ELA and 93% for math while New York and Massachusetts tested a significantly lower proportion of its standards for both subjects (384). As long as the standards tested are random and unpredictable, teaching to the test is not promoted. However, data suggests that in fact, the standards tested *were* predictable. The bottom row of Table 2 reveals that a small proportion of standards are overrepresented and that the “four most highly assessed standards account for over half of all test points between 2006 and 2009 in New York”, with the number of standards being 8 for Massachusetts and 12 for Texas (384). Misalignment between standards and exams, particularly the overrepresentation of certain standards on exams, is present in states through varying

degrees. Texas, by testing the vast majority of its standards, mitigated the predictability of which standards appear on exams, while New York and Massachusetts tested fewer standards which are vastly overrepresented.

Such analysis of alignment reveals the root mechanism of teaching to the test. Teachers are under increasing stress from the standards-based accountability system to improve student achievement, and misalignment means that certain test objects are overrepresented on the state assessments. Therefore, teachers often resort to focusing on these overrepresented items rather than a wholistic curriculum teaching, thus leading to the flourishing of teaching to the test in the NCLB era. Jennings and Bearak give two different ways that educators can respond to this misalignment: “omit content that is never tested on the state test” or “focus on content that predictably composes a large fraction of the state test” (385). Perhaps unsurprisingly, students performed better on the overrepresented standards in all cases except Texas math, which makes sense considering Texas tested most of its standards, thus decreasing predictability for educators. If test items are predictable, teachers tend to narrow their curricula to emphasize these test subjects (385). Further backing this claim is the observation that when a new standard not represented in the previous year is tested for, students performed worse. This can again be explained to teachers focusing their teaching to highly tested standards.

An underlying assumption made above is that due to the pressure of SBA’s harsh sanctions, teachers are driven to teach to the test. Is this an assumption without proof? Research shows that accountability systems affected teachers’ mental health even before NCLB’s passing. In “Put to the Test: The Effect of External Testing on Teachers”, Mary Lee Smith states that the “publication of test scores produces feelings of shame,

embarrassment, guilt, and anger in teachers” (9). Both the administrators and the public view test scores as an evaluation on teaching performance, thus below average test scores call into question the teaching ability of educators regardless of extenuating factors such as socioeconomic status. Even teachers of well-performing students are under increased anxiety and pressure due to administrators’ expectations (9). What do teachers do when faced with this increased pressure? They resort to teaching to the test as explained in the paragraph above. One teacher was quoted as saying, “I wanted to keep my literature program, and I knew if my scores were low, they would make us go back to the basal, so I drilled them with *Scoring High* worksheets [that match the objectives and formats of the ITBS]” (Smith 9). By formatting her curriculum to match questions on the standardized tests to artificially boost test scores, this teacher utilized teaching to the test. Although these are anecdotal evidence, not empirical data, it nonetheless reveals the pressure that SBA causes on teachers. There were also cases in which test anxiety caused students to perform badly on the standardized assessment, even if they performed strongly in classroom tests (Hamilton 121). The sense of powerlessness that teachers may experience in these circumstances may also drive them to teach to the test.

With the passing of NCLB, the accountability system was expanded, causing further pressure and stress, and subsequently driving teaching to the test. Monica Battley-Fabre found in “A Study Examining Teachers’ Perception of State High Stakes Standardized Testing Programs: Their Effects on Teacher Preparation, Performance and Instruction” that across all grade levels, teachers were under significant pressure to raise test scores with teachers experiencing extremely high pressure from pressure from principals (80%) and administrators (75%) (Battley-Fabre 48). This pressure led to 58.3%



of teachers surveyed agreeing that their curriculum was strongly influenced by state tests as outlined in Table 3 below. In particular, 65% responded “yes” and 28.3% responded “sometimes” when asked if they taught to the test more than they normally would. Constructed test items and test preparation were vastly used by a vast majority of teachers (85%), and 93.3% reported frequently covering state items in their instructions (45). Although instruction and standards should ideally be aligned, the increased use of test items and test preparation right before the administering of tests indicates teaching to the test. In the face of these studies, it is evident that accountability system by NCLB promotes teaching to the test (SBA → Increased pressure → Teaching to the test).

Item	Available Range		
NOTE: Numbers in Percentages – TOTAL POPULATION 60			
To what extent do the following factors influence your instructional planning?	Strongly Influence	Somewhat Influence	No Influence
Influence of recent state tests	58.3	41.67	.03
Influence of overall school results	61.7	37	1.3
Influence of last year's performance scores	58.3	41.67	.03
How have state test results affected your instruction?	Yes	Sometimes	Never
I teach to the state test(s) more than I normally would	65	28.3	6.7
I omit certain information because there is not enough time to fit it in because of state test(s)	61.7	23.3	15
I do not do certain things that look interesting or beneficial for students unless they are on the state test	33.3	40	26.7
I do not do anything differently because of the state tests	8.7	44.6	46.7
How frequently are the following practices used in your classroom	Frequently	Sometimes	Never
Constructed response item	85	11.7	3.3
Term projects	10	61.7	8.3
Performance items	65	28.3	6.7
Test preparation	85	13.7	1.3
Covering state test topics	93.3	6.7	0
How much attention are you able to give to the following areas in your classroom	Often	Sometimes	Rarely
Attention given to higher order thinking skills	75	21.7	3.3
Attention given to problem solving skills	71.7	23.3	5
Attention given to factual knowledge	77	21.7	1.3
Attention given to topics not on state test	20	37	43
Attention given to fine and performing arts	16.7	26.7	56.6
Attention given to enrichment	28	37	35

Table 3. Influence of High-Stakes Testing on Curricular Instruction (Battley-Farbe 45).

Why exactly is teaching to the test bad? Yes, it may produce biased test results, but does it worsen the education that students receive? One of the greatest concerns of teaching to the test is its promotion of *curriculum narrowing*. Richard Rothstein and Rebecca Jacobsen explain in “The Goals of Education” that “by basing sanctions solely on math and reading scores, [NCLB] creates incentives to limit – or in some cases to eliminate entirely – time spent on other important curricular objectives” (264). In other words, under pressure from NCLB’s accountability system, educators and administrators target instruction to federally tested subjects, leaving non-tested subjects neglected in its wake. This makes sense because if an educator is under pressure to raise their students’ test scores, they would focus on tested subjects such as reading while neglecting classes such as art. While this may raise scores, it begs the question: are students being deprived of a well-rounded education? One former teacher seems to think so:

From my experience of being an elementary school teacher at a low-performing urban school in Los Angeles, I can say that the pressure became so intense that we had to show how every single lesson we taught connected to a standard that was going to be tested. This meant that art, music, and even science and social studies were not a priority and were hardly ever taught. We were forced to spend ninety percent of the instructional time on reading and math. This made teaching boring for me and was a huge part of why I decided to leave the profession. (Rothstein and Jacobsen 265)

This anecdote reveals that due to SBA’s harsh sanctions, teachers were pressured into teaching to the test. Empirical data backs up this teacher’s claim as by 2007, “71% of districts had reduced instructional time in elementary schools in at least one subject to make more time for English language arts or math (CEP 8). Clearly, curriculum narrowing is not a one-off incident only affecting some schools.

Teachers also increased teaching time for ELA and math (two widely tested subjects on standardized exams) at the cost of subjects such as social studies. A study done

by the National Center for Education Statistics (NCES) titled “Changes in Instructional Hours in Four Subjects by Public School Teachers of Grades 1 Through 4” states that from 1988 to 2004, instructional time for social studies and science decreased, even though *total* instructional time increased over the same period (NCES 2). This can only be attributed to the over-representation of math and English in curricula. On average, instructions for those two subjects decreased by around 20 minutes per week, and although this may seem insignificant, it leads to a deficit of numerous hours over the course of the school year. The same data reveals that from 2003-2004, the percentage of the curriculum dedicated for these 2 “non-essential subjects” were the lowest since 1987. This strongly suggests that NCLB’s strict accountability system is to blame.

Some may argue that this change in instructional hours wasn’t a direct consequence of NCLB. After all, a lot has changed in America from 1988 to 2004, and NCLB could be a confounding variable, not a cause of curriculum narrowing. While this is a valid concern, a study done by the Center of Education Policy (CEP) puts this concern to rest. CEP is an organization that conducted a comprehensive review of curriculum changes under NCLB from 2002-2007. Using a nationally representative and random sample of 491 school districts, researchers were able to determine how the label of “restructuring status”, a disciplinary action under NCLB’s accountability system as explained in Chapter 2, changed the distribution of instructional hours. Table 4 below shows that ELA and math instructional time increased when schools were identified by SBA, backing up the NCES study, while also decreasing instructional time in social studies, science, music, etc., subjects that are not given priority under state tests. These data are significant because while the NCES data showed changes in instructional time *during* the passing of NCLB,

the CEP study directly relates the standards-based accountability system (i.e. restructuring status) to curriculum narrowing. Without the SBA, the “non-essential” subjects would not have taken such a hit. Furthermore, NCLB promoted curriculum narrowing to a significantly further degree than previous accountability systems. As noted in Chapter 2, accountability systems have been in place since at least as far back as Improving America’s Schools Act of 1994. However, “the exclusive emphasis of NCLB on basic academic outcomes is entirely new... Holding schools accountable only for math and reading is an extreme position, which rarely has enjoyed significant support” (Rothstein and Jacobsen 270). In this current system, teachers have no incentives to *not* excessively drill math and English. After all, if that method of teaching brings results and keeps schools out of restructuring status, why would you have a balanced and diverse curriculum? These are questions that must be considered when devising reforms to the NCLB system.

Subject Area	Number of Minutes Per Week		
	Total (All districts)	Districts with No Identified Schools*	Districts with At Least One Identified School*
English language arts	503	483	568
Math	323	320	332
Social studies	178	181	167
Science	178	181	169
Art and music	110	113	97
Physical education	105	106	103
Lunch	142	141	147
Recess	133	134	129

Table reads: In school year 2006-07, districts devoted an average of 503 minutes per week to English language arts in elementary schools. Districts with no schools identified for improvement, corrective action, or restructuring devoted an average of 483 minutes per week to ELA at the elementary level, while districts with at least one identified school devoted an average of 568 minutes to elementary ELA.

\*Identified schools include those identified for improvement, corrective action, or restructuring under the No Child Left Behind Act. Apparent differences between districts with no identified schools and districts with at least one identified schools are not statistically significant, except in English language arts.

Source: Center on Education Policy, February 2007, District Survey, item 18 (tables IT-1A & IT-1B).

Table 4. Instructional Hours and School Improvement Status (CEP).

The neglect of the fine arts has a plethora of consequences as outlined by Claus von Zastrow and Helen Janc in *Academic Atrophy: The Condition of the Liberal Arts in America's Public Schools*. Their finding was two-fold: first, there were increases in instructional time for reading, writing, and mathematics (7). This backs the claims of the NCES and CEP as explained previously. Second, they found that 25% of principals reported decreases in instruction for arts with 33% anticipating decreases in the future. Although these questions were broad and perception-based, they nonetheless reveal how educators view the curriculum narrowing situation. But is this a bad thing? After all, the professions deemed most important to society, i.e. doctors and scientists, are given a STEM education. Zastrow and Janc argue that a liberal arts education fosters an “understanding

of what it means to be human, an understanding that transcends limiting conceptions of occupation, social class, race, or nationality” and makes “active and ethical participants in the world” (11). This cannot be conveyed through a math or science-heavy curriculum. Furthermore, a broad liberal arts education allows one to be adaptable in times of change which “purely technical skills” are doomed to irrelevance before long (11). The study provides several recommendations to promote the liberal arts education, notably to extend NCLB’s standards-based accountability to *all* liberal arts courses (29). The inclusion of subjects such as art or social studies will help mitigate curriculum narrowing.

The erosion of the liberal arts may even threaten the maintenance of democracy. In *Not for Profit: Why Democracy Needs the Humanities*, Martha C. Nussbaum labels this erosion as the “Silent Crisis” in which the humanities are being cut away in favor of subjects deemed more useful. She argues that “thirsty for national profit, nations, and their systems of education, are heedlessly discarding skills that are needed to keep democracies alive” (Nussbaum 2). A liberal arts education strengthens a child’s emotional and imaginative resources, allowing him or her to understand both themselves and others. This leads to generations of “useful machines” because a “technical and factual education” cannot cultivate independently thinking citizens (102). Ultimately, Nussbaum argues, the future of democracy is at stake. The importance of the arts stretches all the way to childhood; for one to understand the complex world, they must have “narrative imagination”, defined as the ability to think in the shoes of the other (Nussbaum 95-96). Because a thorough liberal arts education is crucial to the development of this narrative imagination, the potential impact of curriculum narrowing is clear.

The CBE study introduces a new concept: the “Principal’s Dilemma”. Principals had to choose between allocating more time to raise math/ELA scores and still having adequate time for liberal arts. In essence, there simply was not enough time in the school day to satisfy NCLB’s performance demands while also maintaining a balanced curriculum. Some principals tried to innovate by incorporating mathematics into other classes such as social studies, thus increasing the teaching time of math without cutting other classes, but “without the help of policies and programs that support the liberal arts, principals may prove unable to prevent curricular erosion” (Zastrow and Janc 13). Along with principals, teachers were also plagued by feelings of “too much to do with too little time”. In Donna Newberg-Long’s “Narrowing of Curriculum: Teaching in an Age of Accountability: A Phenomenological Study”, in-depth interviews were conducted with three teachers who taught at the elementary level and had extensive experience teaching in the high-stakes testing era. The purpose of the study included “teacher perception of curriculum narrowing” and “teacher perceptions of how instruction may have altered during their years as a teacher” (Dana Newberg-Long 73-74). All three teachers replied that due to the high-stakes accountability system, district administrators provided rigid curricula in reading and math, thus they did not have the freedom to tailor their teaching to the students’ needs. Social studies took an especially big hit and was regularly shortened to focus on reading, writing, and math. What is the benefit of social studies? Dana Newberg-Long states that “social studies lend themselves to higher level thinking and strategies that support learning for all students, especially Second Language Learners, like visuals and developing increased vocabulary and background knowledge” and “are not as important as they once were with the stress of raising student achievement in tested subjects” (94-95). Furthermore, all 3



teachers state that P.E., art, music, library, and computer science were taught alternatively once a week, a dramatic decrease from pre-NCLB. In short, these interviews emphasize the rigidity of the curriculum to emphasize math and ELA while also cutting subjects such as social studies, all due to the high-stakes accountability system introduced under NCLB.

In response to accusations of curriculum narrowing, some schools responded that they are not cutting out the liberal arts but rather deferring it. According to Craig D. Jerald in “The Hidden Cost of Curriculum Narrowing,” some educators believe that by emphasizing math and English early in the curriculum, students will build the foundation needed to truly understand science and social studies later in their academic career (Jerald 2). At first glance, this seems like a fair argument; how can one read a biology textbook without possessing the adequate reading skills? Although this argument has some merit, learning to read does not necessarily lead to strong comprehension skills. For full understanding of the text, the reader needs to possess “background knowledge about a topic to understand what the text is saying”, or else only partial comprehension occurs (2). Daniel T. Willingham in “How Knowledge Helps” states that authors assume that readers have some level of background knowledge, therefore when this knowledge is not possessed, the reader becomes confused. In one study, poor readers knowledgeable about baseball comprehended a baseball-related text better than good readers who were not knowledgeable about baseball, emphasizing the importance of background knowledge (Willingham 2006). When one is knowledgeable in a given subject, they can devote more effort to visualizing and rationalizing the text rather than trying to comprehend it. Therefore, curriculum narrowing in elementary school leads to inadequate vocabulary and background knowledge in students, leading to reduced reading comprehension in subjects such as

history and science later in their education. With poor reading comprehension, children can fall behind others, and “because background knowledge... is necessary to learn new vocabulary words and information related to that subject, students who start out without background knowledge will all further and further behind” (Jerald 4). It is thus imperative that curriculum narrowing does not occur, especially in early elementary school.

However, the data presented on curriculum narrowing must be viewed through a lens of skepticism. The CBE and Donna Newberg-Long’s studies provide evidence that the *perception* of curriculum narrowing existed. In other words, teachers and principals believed that they are de-emphasizing the liberal arts in favor of STEM/ELA, but this does not necessarily mean that curriculum narrowing occurred. After all, one’s perception of a situation and the actual situation itself can differ significantly. One example is this is the wording of the Zastrow survey. By asking principals, “Did instructional times decrease for arts?”, it introduces some ambiguity because principals can answer the question based on their perception, not actual data. If instead the principals used data on the number of instructional hours to answer the question, then the study would be better rooted in data. Therefore, while data from the Zastrow and Newberg-Long are important, researchers should focus on studies that provide actual data (e.g. NCES and CEP) for proof of curriculum narrowing.

The extent of curriculum narrowing may also be exaggerated. In fact, the hysteria after the passing of NCLB and its supposed “killing of the liberal arts” was overkill and not based on facts. Craig D. Jerald uses the word “erosion” to describe the state of the liberal arts because although more emphasis is being placed on math and English, the arts are not “dying” as so many proclaim (Jerald 2). Although Nussbaum’s arguments of the

“death of democracy” are compelling, it is perhaps farfetched. Undoubtedly, curriculum narrowing is a real consequence of the standards-based accountability system; the CEP and NCES studies provide evidence of this. On the other hand, studies do not suggest that classes such as science or social studies will be cut entirely from the curricula, and trends in instructional data do not suggest this either (NCES 2). The belief that the liberal arts are in imminent danger of death is an over-exaggeration, and such fearmongering is unnecessary when examining the data.

This chapter outlines two of the biggest concerns of No Child left Behind Act’s standards-based accountability system: teaching to the test and curriculum narrowing. Due to Standards-Based Accountability’s harsh sanctions for schools labeled for improvement, teachers and administrators were pressured to tailor their curricula to test items and subjects represented on standardized exams, known under the umbrella term of *teaching to the test*. This was facilitated by misalignment between standards and exams, meaning exams were not equally representative of standards and thus easy to predict by educators. In consequence, certain subjects were over-represented in the curricula, notably ELA and math, while others like social studies and science were relatively neglected. This is known as *curriculum narrowing*. This can deprive students of a valuable liberal arts education which may stunt the development of reading comprehension skills and also interfere with the cultivation of well-rounded citizens, necessary for a well-functioning democracy. However, it is important to remember that much of the studies done focus on the *perception* of curriculum narrowing by educators, and the extent of the curriculum narrowing, though significant, do not suggest the “death of the liberal arts” as some believe.

## CHAPTER FOUR

### Increased Fear and Pressure

Due to NCLB's strict accountability system, more emphasis is placed on standardized test scores than ever before. It is therefore no surprise that fear and anxiety have increased for both teachers, who face consequences for their performance, and students, who are pressured to do well on these tests. Moshe Zeidner explains in *Test Anxiety: The State of the Art* that test anxiety has been flourishing due to the significant long-term impacts of one's standardized test scores. Test anxiety can affect anyone who takes standardized tests, even if it is for the military or the private sector. However, the frequency and importance of standardized exams in schools make children very susceptible to experiencing this anxiety (4). The claim that NCLB harms the mental health of students is serious and will thus be explored further in this chapter.

Zeidner defines test anxiety as the "set of phenomenological, physiological, and behavioral responses that accompany concern about possible negative consequences or failure on an exam or similar evaluative situation" (17). Those who are especially susceptible to stress view tests as threats, leading to lower self-esteem and reduced self-efficacy. This is usually triggered when one believes their "intellectual, motivational, and social capabilities" are insufficient for the demands of the test (18). The level of anxiety experienced is also highly individualistic, with levels of anxiety experienced differing among students. It is important to note that one does not need to have debilitating stress to be considered to have test anxiety; even moderate anxiety during tests is test anxiety.

Patricia A. Lowe et al. in “The Test Anxiety Inventory for Children and Adolescents (TAICA): Examination of the Psychometric Properties of a New Multidimensional Measure of Test Anxiety Among Elementary and Secondary School Students” propose a new model for test anxiety in which the perceived threat (standardized test) affects a test taker’s behavior, cognition, and physiology, known under an umbrella term as “test anxiety” (218). Refer to Figure 3 below. One can react to the perceived threat in two ways: task-relevant behavior (focusing) or task-irrelevant behavior (skimming the text). Physiological symptoms in test anxiety are consistent include sweaty palms and rapid heart rate. Cognitive effects include worry and *cognitive obstruction*, the disruption of one’s “ability to organize his or her thoughts or to concentrate on the task at hand” (Lowe et al. 218). These effects are enough to hinder one’s performance on the standardized test, especially when they respond in a task-irrelevant manner.

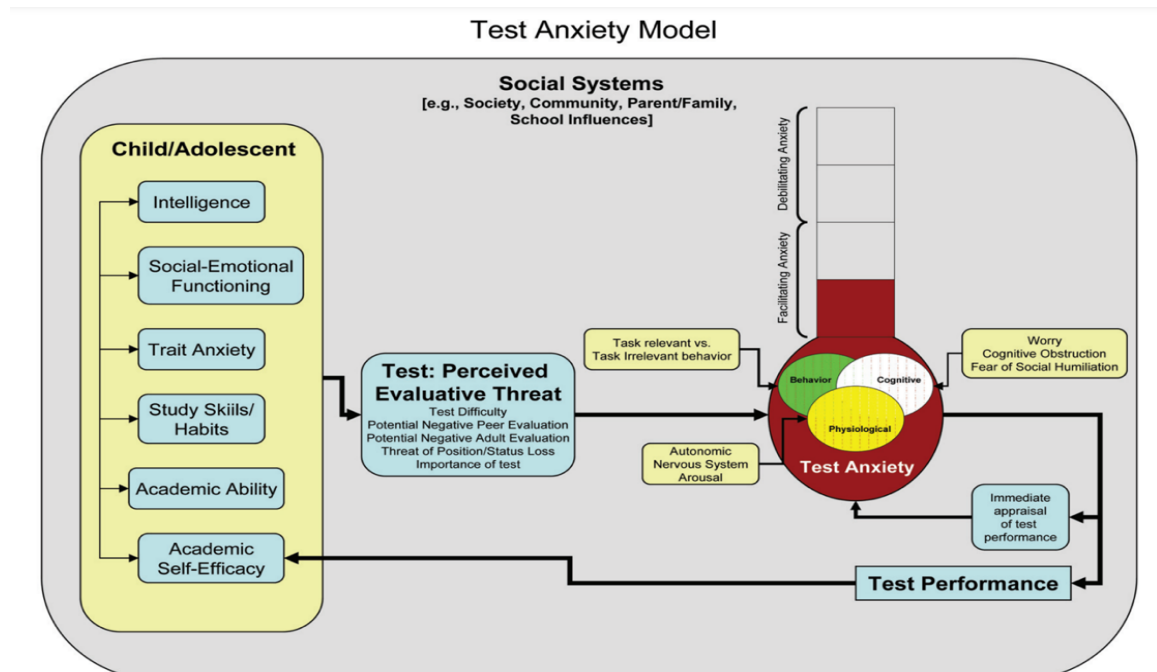


Fig. 3. Test Anxiety Model (Lowe et al. 218).

Surprisingly, there is no clear consensus on the prevalence of test anxiety. Some studies indicate that 25-30% of American students are affected by “debilitating stress” in evaluative conditions, with another Canadian study putting the figure at 22% (Zeidner 7-8). Natasha K. Segool et al. in “Heightened Test Anxiety Among Young Children: Elementary School Students’ Anxious Responses to High-Stakes Testing” back up Zeidner’s claim and estimates the prevalence at 10-30%, although the prevalence of “clinically significant impairment” is likely closer to 10% (490). Furthermore, test anxiety may affect African Americans to a greater degree, up to 41% (Segool et al. 490). Although these studies vary in their results, they confirm the belief that test anxiety affects a measurable number of children. Although some level of stress may enhance performance, stress deemed at an excessive level will harm likely harm performance. The point of standardized tests is to measure the proficiency of students in certain subjects. If the students, ridden with anxiety, perform worse on these tests than their actual abilities, how predictive can these tests really be? More research must be done on the prevalence of test anxiety, particularly in the NCLB era, but the implications are obvious.

A highly competitive and evaluative classroom may increase the incidence of test anxiety. In one study, students who perceived their school to be evaluative and competitive had greater test anxiety (Zeidner 162). In contrast, those with an internal comparison point (comparing themselves to their own previous performance) had lower anxiety. Because of NCLB implemented yearly testing with significant ramifications, it is reasonable to assume that it led to a more competitive, and undoubtedly a more evaluative, classroom environment. In “Third Through Sixth Graders’ perceptions of High-Stakes Testing”, Cheri Foster Triplett and Mary Alice Barksdale asked elementary school students to draw

a picture representing their high-stakes testing experience. They found that both NCLB-mandated tests caused significant worry and anxiety. When asked by researchers, the students responded that the “consequences of failing the test” caused this anxiety (Triplett and Barksdale 225). These perceived consequences ranged from summer school to a future in fast food. It is unsurprising that very few students had positive experiences given the high-stress testing environment. In today’s climate of increased mental health awareness, the ethics of subjecting young children to such anxiety must be considered.

Although it is perfectly conceivable that NCLB is responsible for increased test anxiety, one may argue that normal classroom test may cause test anxiety at comparable rates to high-stakes exams. Therefore, to avoid making unsubstantiated assumptions, Natasha K. Segool et al. studied the specific impact of NCLB. They used CTAS and BASC-2-TA, self-reported questionnaires that determined the level of test anxiety. Students were then placed into three groups: high anxiety, moderate anxiety, and low anxiety (Segool et al. 494). Anxiety scores were collected twice, first after NCLB-mandated exams and then after normal classroom tests, crucially allowing for comparison. Previous studies such as Triplett and Barksdale’s mentioned above only measured anxiety during high-stakes exams, meaning such a comparison to classroom exams could be made. Segool et al. found that for both scales, test anxiety for high-stakes exams were greater than for classroom exams at a statistically significant level(493). Surprisingly, the percentage of students labeled “high anxiety” stayed approximately constant at around 10% from NCLB to classroom tests. The main difference in anxiety scores came from the “moderate anxiety” and “low anxiety” groups. As illustrated in Figure 4, “more students reported moderate test anxiety (59% vs. 44%) and fewer students reported low test anxiety (32% vs.

45%) about NCLB testing versus classroom testing” (Segool et al. 496). This suggests that the students who experience debilitating stress will feel this stress regardless of the testing condition. Indeed, Beth Ann Fulton in “The Relationship Between Test Anxiety and Standardized Test Scores” suggests that test anxiety may be, in part, hereditary and rooted in biological roots (22). She suggests that to understand a student’s test anxiety, knowing his or her family history could be useful.

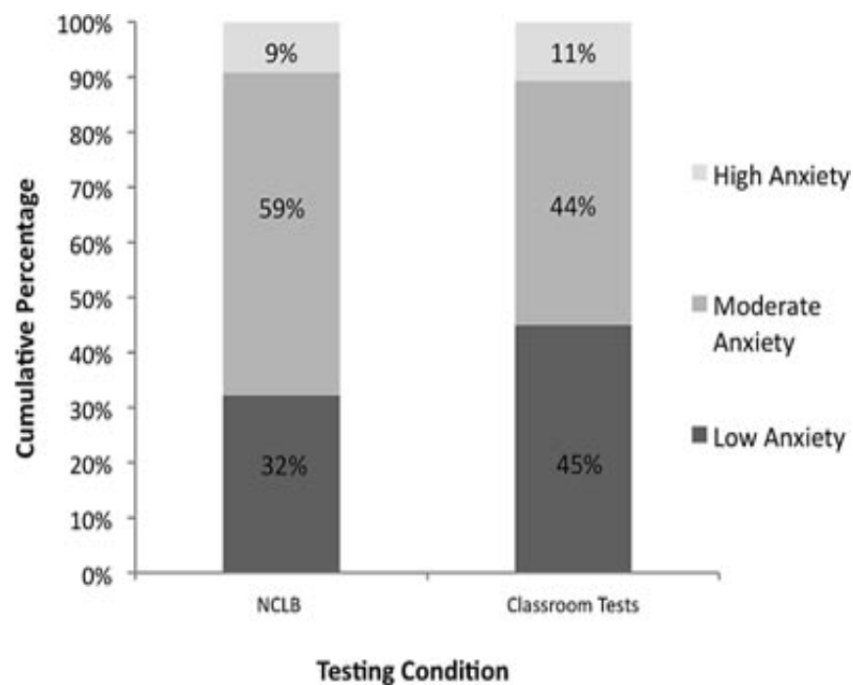


Fig. 4. Test Anxiety between Classroom and NCLB-mandated Tests (Segool et al. 495).

It is important to note that there are multiple variables that affect test anxiety. Along with the biological aspect mentioned above, Fulton explains that gender, age, socioeconomic status, and race are also correlated to one’s test anxiety (22). Females are more likely to have test anxiety compared to males, possibly because they are “more uncomfortable and self-conscious in testing situations than men” (Fulton 22). However, if



females are more likely to report test anxiety, this may skew results. This is backed up by Segool et al. who found that females scored higher on both the CTAS and BASC-2-TA at a small but statistically significant level (494). Furthermore, test anxiety typically increases with age, likely due to harder material increasing pressure from teachers and parents (Fulton 23-24). This is not to say that elementary school children do not experience anxiety all, just that they experience it to a lower degree. The degree to which race affects test anxiety is questionable; while some claim that certain races (notably Asians) experience more anxiety due to cultural differences, Segool et al. find no correlation between ethnicity and test anxiety (Segool 494). This highlights the complexity of test anxiety and the multitude of factors that may lead to its development.

However, researchers are not unanimous in linking high-stakes exams to test anxiety. In the study “Impact of Accountability and School Testing on Students: Is There Evidence of Anxiety?”, Sean W. Mulvenon et al. measured test anxiety after the fall 2000 Stanford Achievement Test and the 2000 Fourth Grade Benchmark Examination. Similarly to Segool, they utilized a questionnaire to determine test anxiety. Surprisingly, researchers found that most students did not experience stress under the standardized tests, with some even having a positive experience (15). In one way, this study backs up Segool in that standardized tests do not increase the number of “high anxiety” students. However, to say that students did not experience test anxiety under standardized exams goes against the other studies mentioned so far. These results may be due to several forms of bias. First, the students had to get consent from their parents to participate in the study, meaning the sample was not truly random (13). It is possible that parents of high-anxiety students declined consent, perhaps in fear of causing more anxiety, skewing the results of the study.

Second, only a single southern school district was studied (12). Given the small sample size, one cannot extrapolate the results to all American school districts. This is not a problem limited to Mulvenon; Segool's study only examined three elementary schools in the Midwest (Segool et al. 491). This points to a need for a more comprehensive study for test anxiety. Third, only 5<sup>th</sup> graders were studied. Test anxiety can affect students from young children to adolescents, and this wide range in age was not accounted for. Segool's study was better by studying 3<sup>rd</sup> to 5<sup>th</sup> grade (Segool et al. 491). Most significantly, the study was completed *before* the passing of No Child Left Behind Act, meaning that its results cannot automatically be applied to the NCLB era. Given all this, it would be more prudent to trust Segool's conclusion that NCLB does indeed raise rates of moderate test anxiety among students, although the level of high anxiety students likely remained constant.

Along with students, studies show that teachers are also experiencing anxiety from high-stakes exams. Anecdotes from one teacher reveals that the release of test scores caused them anxiety, especially because of the harsh sanctions under NCLB (Smith 9). Interestingly, some teachers felt guilt over subjecting their students to standardized exams, even describing them as "cruel and unusual punishment" (Smith 9). In other words, the students' test anxiety drove anxiety in teachers. Many teachers also worried that they had inadequately prepared students for the test, and in response, may "change their instructional patterns to focus on test preparation" (Segool et al. 496). Curriculum tests are made by teachers themselves to assess the curriculum material, therefore teachers would not feel the same anxiety of "what if I taught the wrong material?". This issue goes back to the alignment between assessment and instruction, which is often difficult considering

NCLB's wide breadth and depth of standards as well as insufficient horizontal misalignment. Thus, it appears that NCLB, with its promotion of high-stakes exams, harms the mental health of both students and teachers by increasing rates of stress and anxiety.

Are there disadvantages to test anxiety beyond harming the mental health of students? Many assume that test anxiety leads to worse scores, and logically, this makes sense. When one imagines test anxiety, the image of sweaty hands, a racing heart, and the panic of their mind going blank probably comes to mind. In her study, Fulton measured test anxiety in two ways. First, she used a questionnaire not unlike those used by Segool and Mulvenon as a subjective measure of test anxiety. Surprisingly, the correlation between anxiety score and test results was  $r=-0.1603$ , indicating an insignificant relationship (Fulton 52). Second, she took the pulse of students during the exam as an objective measure of test anxiety with the assumption that anxiety raised one's heart rate, getting a moderate negative relationship. This means that as heart rate increased, the student's test scores decreased (52). This study suggests that questionnaires, which are subjective in nature, do not necessarily correlate to physical symptoms of test anxiety. It is possible that students underestimate or downplay their anxiety, skewing questionnaire data, while objective tests such as heart rate show the "true" amount of test anxiety. Future studies could look at perspiration, rapid breathing or cortisol levels for other physical measures of test anxiety.

Test anxiety does not necessarily equate to worse performance. According to Zeidner, test anxiety can be one of two types: *facilitating* or *debilitating*. Facilitating anxiety enhances test performance while debilitating anxiety inhibits performance, forming a continuum with some experiencing mostly facilitating anxiety, some experiencing mostly debilitating, and others lying somewhere in the middle (208). Furthermore, many factors

beyond test anxiety are involved in academic performance, countering the myth that *everyone* with test anxiety performs worse. After all, studies show that “many individuals who are test-anxious nevertheless perform well on cognitive tests” (Zeidner 235). In general, though, research indicates that there is a moderate inverse relationship between anxiety and test performance (234). In other words, as levels of anxiety increases, students tend to perform worse. The association is not very strong with studies ( $r = -0.20$ ) although still statistically significant (Zeidner 235). In conclusion, test anxiety does not automatically correlate to worse performance as commonly thought. However, those who experience *debilitating stress* may perform better under stress. Given this, the general trend does indicate that test anxiety is a net disadvantage.

This begs the question: what can be done to reduce test anxiety? Standardized testing has become so entrenched in the education system that getting rid of all anxiety-inducing exams is not realistic. In a comprehensive review of various test-anxiety interventions, Nathaniel von der Embse, Justin Barterian, and Natasha Segool found in “Test Anxiety Interventions for Children and Adolescents: A Systematic Review of Treatment Studies from 2000-2010” that one promising method of intervention is Cognitive-Behavioral Therapy (CBT) which has been shown to significantly decrease test anxiety and even increase grade point averages (von der Embse et al. 67). Though these results are fantastic, therapy may not be financially feasible when applying to a large group of students. Interventions that can be applied to a wider population of students usually use “cognitive or behavioral theory” which includes relaxation and biofeedback techniques (67). In fact, every effective intervention reviewed by the authors used these theories. Several interventions will be reviewed in the following paragraphs.

Teachers can create a supportive classroom environment and implement techniques to manage test anxiety. Alicia Stapp and Ashley Lambert explain in “The Impact of Mindfulness-Based Yoga Interventions on Fifth-Grade Students' Perceived Anxiety and Stress” that mindfulness, the “intentional focus and awareness on one’s present environment”, has gained traction to reduce anxiety in the NCLB era (471). Mindfulness “fosters thoughtful response rather than immediate actions” to stressful situations (472). Stapp and Lambert studied fifth grade students’ anxiety and stress pre- and post-intervention, measured through questionnaires and interviews with teachers. The intervention involved breathing exercises, slow counting, and various yoga stretches (Stapp and Lambert 474). Interestingly, intervention produced different results by gender. Male students showed reduced anxiety and stress post-intervention while female students’ anxiety rose by 5%. Furthermore, intervention reduced anxiety the greatest in remedial students and least in advanced students (475-476). Mindfulness intervention does not seem to be the universal solution for test anxiety; males in remedial courses would benefit the most while females in advanced courses may even be harmed, shown by their rise in stress levels post-intervention (476). However, anecdotal evidence from teachers shows that as a whole, “there was an ‘obvious’ decrease in anxiety and stress, ‘especially around test days’” (Stapp and Lambert 78). This included calmer behavior and less concern over grades.

More extensive interventions beyond simple yoga stretches show greater efficacy in reducing test anxiety. Jeanne M. Donato in “Reducing test anxiety and improving academic performance in fourth grade students: Exploring an intervention” studies the efficacy of the *TestEdge* intervention on test anxiety. *TestEdge* is an eight-lesson curriculum that includes “deep attitudinal breathing, emotional self-management tools such

as Freeze-Frame, Neutral, specialized music...and a three minute computerized biofeedback game” (Donato 27). Instructors of fourth grade students were provided with a manual, workbook, and photo cards giving specific instructions on TestEdge. This multi-lesson intervention is notably far more comprehensive than Stapp and Lambert’s yoga stretches. The students were randomly assigned into control and experimental groups. Surprisingly, there was no statistically significant difference in pretest mathematics scores between the groups, nor the posttest scores. This indicates that the *TestEdge* intervention did not change the academic performance of the students. However, the experimental group showed a significant increase in *coherence scores* (Donato 215). According to Donato, “*Coherence Scores* reflect the level of stress a person experiences in real time” using a biofeedback program mentioned above which indicates that the *TestEdge* intervention significantly decreased student stress compared to the control (13). This contrasts with CTAS data which showed that levels of test anxiety were also not statistically different between the control and experimental group (Donato 215). Thus, the *TestEdge intervention* provides a promising way to reduce stress in students but should not be relied on to change the students’ academic performance.

The No Child Left Behind Act and its accountability system have raised the stakes for standardized testing at an unprecedented level. Faced with pressure from teachers and administrators, a plethora of students are at risk of experiencing test anxiety, even affecting test performance for some who are particularly susceptible. In lieu of this, schools have an obligation to implement intervention techniques to control students’ test anxiety, with the most effective interventions using Cognitive and Behavioral Theory including biofeedback systems, relaxation techniques, and Cognitive-Behavioral Therapy. However, even with

test anxiety affecting upwards of 40% of students, research on test anxiety has been declining since the 1970's (von der Embse et al. 57). This points to a need for further research on effective interventions, and most importantly, implementation of these programs into school curricula. To subject students to an increasingly stressful testing environment without giving them the tools to deal with it is unethical.

## CHAPTER FIVE

### Evaluating SBA and Recommendations for Improvement

The previous four chapters have outlined some of the major concerns of Standards-Based Accountability provision of the No Child Left Behind Act. This provision has attracted lots of criticism from researchers with many accusing SBA of promoting teaching to the test, curriculum narrowing, and test anxiety. Given these host of criticisms, one may ask, “why not abolish standardized testing?” This is indeed a viewpoint shared by some researchers who believe that no amount of reform can save standardized testing. Some alternatives suggested include sampling, or administering the tests to a representative sample of students rather than every student. This invites numerous questions: what exactly does “representative” mean? Is it fair to suggest only some students to these exams? How does one stop schools from testing only their brightest students?

Another alternative suggested is the Portfolio Assessment. According to Ricky Lam in “Understanding Assessment as Learning in Writing Classrooms: The Case of Portfolio Assessment”, a portfolio assessment is a compilation of a student’s work throughout the year. This allows students to showcase their growth at intervals more frequent than NCLB-mandated exams. By compiling works for the portfolio themselves, students must “make informed decisions about the quality of their portfolio works” which is much more interactive than standardized exams (Lam 26). However, major drawbacks to the portfolio assessments remain. Portfolios are a conglomeration of students’ works including essays and projects, thus the question of how to fairly score the portfolios remain. Each student’s



portfolio is different in content from others' which makes direct comparisons difficult. Who scores the portfolios? By nature, portfolios are highly individualistic, therefore a rubric must accurately score a vast array of works. This will no doubt introduce a level of ambiguity in scoring not seen in the current standardized testing scheme. Another criticism, mentions Lam, is the fear that students may not compile a portfolio that accurately marks their growth, instead producing a "promising narrative about their learning in fear of losing marks" (28).

These are questions that must be asked when before blindly calling for the abolishment of standardized tests. To implement a new testing system without adequate research may do more harm than good. Furthermore, for all its criticisms, standardized tests form a very real purpose: assess students' capabilities in an objective manner. Because these assessments are standardized (hence the name), comparisons can be drawn state or nation-wide, and educators are able to see if they are teaching effectively. Standardized tests thus provide invaluable information to the education system. According to "Filling in the Blanks: Putting Standardized Tests to the Test" by Gregory J. Cizek, standardized testing is merely a tool, much like a hammer. A hammer is useful when building, but in the wrong hands, it has the potential to commit murder. However, this potential for misuse does not mean that all hammers should be vilified (Cizek 44). There is a plethora of reasons that standardized tests remain so popular. Compared to alternatives like portfolio assessments, they "provide more information about students, at less cost, and with less student testing time" with the added benefit of a quick "turn-around time for score reporting" (Cizek 44-45). Furthermore, although portfolios can showcase individual achievement well, its unique nature prevents direct comparisons between students or groups (Lam 28).

Voices calling for standardized testing's abolishment may not be as numerous as people may think. According to Richard P. Phelps in *Defending Standardized Testing*, the "testing backlash" that is so often mentioned is not as widespread as one may think. Criticisms that standardized testing erodes democracy and will cause impending doom, stated by Nussbaum, is in fact a minority opinion. Phelps points out in 2001, the positive differential in favor of NCLB's testing requirement for the general public ranged from 48% to 65% (14). If 73% viewed NCLB as favorable and 25% as unfavorable, this would be a 48% differential, meaning that public support was overwhelmingly in favor of NCLB that year. Likely due to the stricter accountability system, however, teachers' support for testing has declined since the 1970's. Therefore, although some erosion in support for standardized testing exists, it is not the backlash that many believe it to be.

The criticisms of NCLB outlined in this paper, namely teaching to the test, curriculum narrowing, and test anxiety, is traced back to the unprecedented standards-based accountability system, as detailed extensively in previous chapters. When schools failed to meet the Adequate Yearly Progress, the accountability system acted as a "punishing system". This can range from providing students with transportation to better-performing schools to the complete restructuring of the school by replacing the school staff (Forte 81). Faced with such pressure, teachers resorted to teaching to the test in attempts to raise student scores. The focus on tested subjects also leads to curriculum narrowing with non-tested subjects such as social studies receiving less focus. Furthermore, the pressure from high-stakes exams increased the incidence of test anxiety, although those who experience "debilitating anxiety" may have remained stable (Segool et al. 496).

To mitigate the negative consequences outlined, the accountability system should be changed to be less punitive. The numerous benefits of standardized testing suggests that reform, not abolishment, of NCLB is the solution. After all, the high-stakes exams mandated under NCLB provide invaluable information to teachers'. While schools at the cusp of making Adequate Yearly Progress showed willingness to improve, those with a "significant score deficit to make up" felt that AYP was an unrealistic goal. Instead of improving the teaching, one school introduced an IB program to lure back high-performing students from magnet schools, lowering the proportion of disadvantaged students in the classroom (Brown and Clift 786). This attempt to change the school demographics is at the very least unethical and not consistent with the goals of NCLB. Thus, educators and administrators may develop learned helplessness due to a seemingly impossible task of making AYP. The pace of SBA may also be too harsh; schools were given 4 years to meet AYP before restructuring which may not be enough to underperforming schools. At McKean Elementary School, a school that failed to meet AYP for 3 consecutive years, panic caused the principal to be replaced yearly, "undermining even promising reforms" (Brown and Clift 787). To punish underperforming schools instead of providing necessary aid goes against all logic. Wang et al. even compares SBA to a yardstick that has been turned into a beating stick, "punishing teachers and students..." (319). When a puppy defecates inside the house, what does one do? Potty train it or spank it? The obvious answer is to train the puppy, to help it so it does not make the same mistake in the future. In the same way, SBA should seek to help struggling schools to make the AYP by decreasing the magnitude of the sanctions and perhaps even allocating more funds to them. With these

changes, educators will be able to focus on improving their curriculum instead of resorting to measures such as teaching to the test.

If the SBA system's punitive measures were too harsh, some may ask: Why have an accountability system at all? If SBA is the root of problems like teaching to the test, shouldn't the whole system be abolished? History has shown us that without a proper accountability system, schools can easily abuse any funding. When the Elementary and Secondary Education Act (ESEA) was passed in 1965, it did not feature any accountability system (Rhodes 33). This led to schools using Title I funds to build swimming pools and buy band uniforms, not to improve education quality as intended (Gamson et al. 15). ESEA was ultimately unsuccessful in closing the achievement gap, in no small part due to the absence of an accountability system. Teachers must also be held accountable for the quality of the education; if they are not, what is the incentive for them to improve their teaching? Like all professions, one should be held responsible for his or her job performance with consequences if the performance is not satisfactory.

Under NCLB, the SBA system mandated states to establish standards for reading (ELA), math, and science. Annual assessments were mandated for ELA and math from grades three through eight, with science being tested only twice in that time span (Hamilton 10). This focus on ELA and math has led to curriculum narrowing with instructional time decreasing for non-tested subjects like art and social studies (CEP 7). This is a major criticism of NCLB's accountability system and one worth examining. This is not to downplay the significance of ELA; undoubtedly, English and math skills provide an invaluable base for students. Without good reading comprehension, how can one properly understand a history textbook. Likewise, without a solid math base, how can one learn

subjects such as physics? Firstly, having background knowledge in a particular subject improves reading comprehension in that subject (Jerald 4). For example, to truly understand the Declaration of Independence, one should know the history behind its passing; good reading skills can only get you so far. Furthermore, the purpose of this paper is not to argue against ELA and math's importance, but rather to draw attention to the decline in "non-essential" subjects. Rothstein and Jacobsen advocate for a more balanced accountability system that include subjects beyond ELA and math (271). Additional subjects included under SBA could be physical fitness, art, and debate. This will reduce the incentive to narrow the curriculum in favor of math and ELA and foster a more balanced curriculum. To ensure that schools are not neglecting the arts, Rothstein and Jacobsen suggest the creation of an inspections team that will "judge not only the quality of school facilities, but also the quality of instruction" (272).

More effort should be placed into aligning the state standards to test material. Misalignment between these two, known as "horizontal misalignment", means that certain test objects are overrepresented on exams and thus easier to predict by educators (Case et al. 9). Horizontal misalignment can also undermine the purpose of the standardized assessment. Ideally, the assessment should test students' proficiency in the standards, which cannot be done effectively if misalignment between the two occurs. Misalignment also allows educators to focus on these test items and thus teach to the test. The extent of the misalignment differs by state with New York only testing 60% and 41% of ELA and math standards respectively (Jennings and Bearak 384). The American Federation of Teachers in 2006 report similar results with alignment only being 60% and 53% in Georgia and Pennsylvania respectively (Hamilton 133). California was found to have 100%

alignment and could serve as a model for other states. Perhaps unsurprisingly, students performed better on standards that were more frequently tested. By bringing the standardized assessments closer to the state standards, teachers will be less likely to focus on overrepresented items, instead teaching the state-mandated standards in a more balanced manner.

Misalignment can also occur between the standards or assessments and the curriculum. In other words, the school curriculum is not representative of what is tested. In California, over 90% of math teachers sampled stated that they aligned their instruction with the standards, but only around 50% stated they also aligned their instruction to the standardized test (Hamilton 101). Responses by ELA teachers were also similar. This would not be an issue if standardized tests were aligned perfectly with standards as intended, which is unfortunately not the case as Jennings and Bearak pointed out. A significant reason for this may be the sheer number and depth of the standards. Many teachers complained that they were simply too numerous to cover in-depth. This can lead to “an excessive reliance on the test rather than the standards to decide what to teach” (Hamilton 133). Reducing the number or depth of the standards would help teachers in the alignment of their curriculum.

Many instructors believed that assessments mandated under NCLB were too challenging. This was especially the case for minorities and non-English speakers. When interviewed, their instructors felt that making AYP for these students was an impossible task (Brown and Clift 785). Because of this sense of helplessness, practices such as teaching to the test and curriculum narrowing, along with the increased rates of test anxiety, resulted under NCLB. Therefore, standardized assessments should feature better

*instructional match*. “Handout 3: Examples of Evidence-Based Interventions” by the National Center on Intensive Intervention (NCII) defines instructional match as the “accurate assessment of a student’s current instructional level and the selection of appropriately matched curricula...” (6). This will ensure that students are taught materials that are adequately challenging for them; if the material or standardized exam is too difficult, students’ learning will be hindered, and they would most likely become very frustrated. Likewise, materials that are not challenging enough for students will lead to boredom and wasted time. NCII states that for proper instructional match, the student’s ability level must be assessed accurately. Students can then be taught at an *instructional level* which allows students to be engaged while also needing some guidance (NCII 7).

Instructional match allows students to learn most effectively. Edwards J. Daly III in “The Effects of Instructional Match and Content Overlap on Generalized Reading Performance” explored the impact of appropriate instructional match on students’ generalized reading performance. When students were taught to read with passages that were adequately challenging (instructional match), their reading speed and accuracy increased. Likewise, when the assessments matched the difficulty and content of the instruction (content overlap), students performed better (Daly III et al. 514). Content overlap contributed significantly to reading skills, with low-overlap exams producing 16.6 % lower accuracy and eight fewer words per minute compared to high-overlap exams (Daly III et al. 514). One drawback of the study is that the participants were all below grade level with some having learning disabilities. Because NCLB encompasses students from a wide range of abilities, the results from this study cannot automatically be applied to NCLB assessments. However, in school districts where meeting AYP seems like an impossible

task due to the assessments' difficulty relative to students' levels, instructional match with the assessments may bring promising results. Perhaps by reducing the depth and number of standards as suggested by Jennings and Bearak, the standardized exams would test students at a level that is appropriately challenging for the struggling students. Teachers should also ensure that the "gifted" and high-achieving students are still adequately challenged and not bored by the curriculum; to sacrifice the education of one group of students is not the solution for NCLB's problems.

By 2010, less than ten years after NCLB's passing, public support for the act had crumbled. Its original goal of all states achieving 100 percent proficiency by 2014 was nowhere near to being realized. On top of that, the standards-based accountability system meant to motivate schools to improve was widely viewed as overly harsh. In *The Every Student Succeeds Act: What It Means for Schools, Systems and States*, Frederick M. Hess and Max Eden state that by 2004, no state was on track to staff all classrooms with highly qualified teachers. In 2007-2008, "35 percent of all schools failed AYP, and this rose to 48 percent in 2011" (Hess and Eden 33). With more and more schools rising through the levels of sanctions, doubts against NCLB increased. States also tried to "cheat" AYP by lowering their proficiency cutoffs, leading to 82 percent of Massachusetts schools failing to meet AYP compared to 16 percent in Kansas (Hess and Eden 33). One of standardized testing's biggest strengths is the ability to allow comparisons between groups; with such ambiguity in proficiency cutoffs, how can schools be truly compared?

The numerous complaints against NCLB culminated in the passing of the Every Student Succeeds Act (ESSA) in 2015. The ESSA was a re-authorization of the ESEA of 1965 and the replacement for NCLB. It passed with overwhelming bipartisan support in



Congress, emphasizing how unpopular NCLB had become by then (Hess and Eden 33). “ESSA and Accountability: Frequently Asked Questions” by the Association for Supervision and Curricular Development (ASCD) outline the major changes under ESSA. States could now test proficiency in subjects beyond math, ELA, and science, with academic and nonacademic factors allowed by the school’s choice. This means that subjects such as art or history could now be tested, reducing fears of curriculum narrowing. Nonacademic measures could be school climate and safety, student engagement, or other measures of the state’s choosing (ASCD 2).

Perhaps the most significant development was the changes to the accountability system. States were given the power to set proficiency goals and implement their own accountability systems (ASCD 1). This is a reversal of the general trend to give more authority to the federal government as explained in Chapter 1. In particular, ESSA sought to change the punitive nature of NCLB’s SBA system. States now must allocate 7% of their Title I funds for “school improvement efforts”, an increase from 5% under NCLB (Hess and Eden 81). This will work to help, not hurt, underperforming schools. Schools also no longer had to approve every single purchase made with Title I funds. Under NCLB, schools had to prove that purchases with Title I funds would not have been made without federal aid which limited how they could spend it. ESSA removes the need to approve every purchase with Title I funds (Hess and Eden 86). With the states now able to set their accountability systems, the federal government can no longer mandate sanctions for underperforming schools. These harsh sanctions placed significant pressure among teachers and educators, therefore ESSA will go a long way in reducing these problems. Another complaint of NCLB’s accountability system was its level-based measure of achievement

(Hess and Eden 80). Whether a school met AYP was determined solely by the proportion of students deemed proficient in math and ELA. Under ESSA, a growth-based measure of achievement was implemented in which students' improvement in scores throughout the years can determine achievement. Hess and Eden state that this is a "more accurate gauge of schools' contributions to student learning" (80), because a level-based system like NCLB is more susceptible to extraneous factors such as socioeconomic status.

ESSA provides flexibility to the accountability system that NCLB did not have. In "Every Student Succeeds Act: Early Observations on State Changes to Accountability Systems", the U.S. Government Accountability Office (GAO) interviewed national stakeholder groups that worked directly with state accountability systems as well as educational agencies and school districts (2). Most stakeholder groups considered ESSA's accountability provisions as a "good balance between flexibility and requirements" (GAO 3). In other words, states had the freedom to design their own accountability systems but were still obligated to provide a good education for students. States must still identify underperforming schools, but unlike NCLB, states can decide which academic factors are considered by the accountability system and the relative weight of each factor (ASCD 81, Hess and Eden 81). Because NCLB accountability provisions were uniform across states, a "one size fits all" approach had to be implemented. This was counterproductive when states can differ significantly in standardized test scores, demographics, and funding. ESSA allows states to implement an accountability system that best serves its own interests.

By Fall 2017, all states submitted their proposals for the accountability system. In "Proposed State Accountability Systems Under the Every Student Succeeds Act: A Summary of Fall 2017 Submissions", David English provides an overview of these

submissions. States took advantage of their added flexibility by including non-academic measures into the accountability system; 74% of states proposed a “college and career readiness measure” and 26% wanted participation in a well-rounded education (English 1). This is more holistic than NCLB’s SBA which only considered scores in tested subjects. Also, states are shifting away from a simple levels-based measure of achievement to one that encourages and measures student progress (English 1). Another significant finding is the inclusion of science and social studies in several states. Texas notably includes science, social studies, and writing in addition to ELA and math into its accountability system (English 3). This broadening of the accountability system will mitigate some of the curriculum narrowing that occurred under NCLB.

The accountability system under ESSA is less punitive, more flexible and more holistic than NCLB’s, which will in theory help instances of teaching to the test, curriculum narrowing, and test anxiety. Teachers can now ideally teach *for* the test, which is teaching that “prepares students to demonstrate their acquired knowledge of a broadly defined subject domain...” (Crocker 9). This will allow students to gain non-superficial knowledge of the state standards that extends beyond test taking skills. Teaching for the test requires comprehensive instruction by teachers which ESSA’s new accountability provisions will encourage compared to NCLB’s. Although some people may decry ESSA’s reversal of the trend of federalization in standardized testing, its novel reforms go a long way in answering some of NCLB’s biggest criticisms.

## CHAPTER SIX

### Conclusion

In conclusion, NCLB did not achieve the widespread improvements in education as originally intended. One of the biggest criticisms of the act is its accountability system. Although schools need to be held accountable of their students' learning, SBA took it to a level that was considered too punitive by many. Faced with the prospect of receiving SBA's sanctions, rates of teaching to the test and curriculum narrowing increased, with studies revealing that instructional time for subjects such as social studies and art were sacrificed for math and ELA. Furthermore, independent studies confirm higher rates of test anxiety in students, some of whom experienced debilitating anxiety that negatively affected their performance on the standardized exam. These are significant issues that cannot go unignored. Identifying what is responsible, which is the harsh accountability system, is the first step in solving these issues.

The plethora of consequences suggests that the accountability system had to be reformed. Some suggestions include making the sanctions less punitive, increasing alignment between standards and assessments, and including more subjects considered under SBA. Under NCLB, many schools, particularly in low socioeconomic areas, complained that AYP was too difficult and felt impossible to meet. Reducing the depth or range of standards will low AYP to be a more attainable goal.

In response to these complaints, President Barack Obama passed the Every Student Succeeds Act in 2014, replacing NCLB. This act incorporated many of the

suggestions listed above, notably by giving states the right to set up their own accountability system. These new accountability systems are more flexible, comprehensive, and less punitive than NCLB's SBA. ESSA bucks the trend of federalization of standardized testing and a significant development of the ongoing story of standardized testing in America.

NCLB's standards-based accountability system had good intentions. Its purpose was to identify and improve underperforming schools, thus "leaving no child behind". Indeed, history shows us that without an accountability system, schools can misuse federal funding and neglect improving student education. However, studies in the past two decades indicate that the overly punitive nature of SBA harmed the educational quality of students across America. As states implement their own accountability systems under ESSA, these issues can hopefully become relics of the past.

## BIBLIOGRAPHY

- Alicia Cooper Stapp and Ashley Berry Lambert. "The Impact of Mindfulness-Based Yoga Interventions on Fifth-Grade Students' Perceived Anxiety and Stress." *International Electronic Journal of Elementary Education*, vol. 12, no. 5, 2020, pp. 471–80. *baylor.primo.exlibrisgroup.com*, <https://doi.org/10.26822/iejee.2020562137>.
- Association for Supervision and Curricular Development. *ESSA and Accountability: Frequently Asked Questions*. [https://files.ascd.org/staticfiles/ascd/pdf/siteASCD/policy/ESSA-Accountability-FAQ\\_May112016.pdf](https://files.ascd.org/staticfiles/ascd/pdf/siteASCD/policy/ESSA-Accountability-FAQ_May112016.pdf).
- Battley-Fabre, Monica. *A Study Examining Teachers' Perception of State High Stakes Standardized Testing Programs: Their Effects on Teacher Preparation, Performance and Instruction*. ProQuest Dissertations Publishing, 2011. *baylor.primo.exlibrisgroup.com*, <https://search.proquest.com/docview/880573000?pq-origsite=primo>.
- Betebenner, Damian W., et al. "The Impact of Performance Level Misclassification on the Accuracy and Precision of Percent at Performance Level Measures." *Journal of Educational Measurement*, vol. 45, no. 2, June 2008, pp. 119–37. DOI.org (Crossref), <https://doi.org/10.1111/j.1745-3984.2007.00056.x>.
- Brown, Abigail B., and Jack W. Clift. "The Unequal Effect of Adequate Yearly Progress: Evidence From School Visits." *American Educational Research Journal*, vol. 47, no. 4, 2010, pp. 774–98.
- Casalaspi, David. "The Making of a 'Legislative Miracle': The Elementary and Secondary Education Act of 1965." *History of Education Quarterly*, vol. 57, no. 2, May 2017, pp. 247–77. EBSCOhost, <https://doi.org/10.1017/heq.2017.4>.
- Case, Betsy J., et al. "Alignment in Educational Assessment." *Pearson Education*, Dec. 2004, pp. 1–13.
- Cizek, Gregory J. *Filling in the Blanks: Putting Standardized Tests to the Test*. 1990.
- Daly III, Edward J., et al. "The Effects of Instructional Match and Content Overlap on Generalized Reading Performance." *Journal of Applied Behavior Analysis*, vol. 29, no. 4, 1996, pp. 507–18. *baylor.primo.exlibrisgroup.com*, <https://doi.org/10.1901/jaba.1996.29-507>.

- Daniel T. Willingham. "How Knowledge Helps | American Federation of Teachers." *American Educator*, vol. 30, no. 1, Spring 2006, <https://www.aft.org/periodical/american-educator/spring-2006/how-knowledge-helps#back1>.
- Dillon, Erin, and Andrew J. Rotherham. "States' Evidence: What It Means to Make 'Adequate Yearly Progress' Under NCLB." *Education Sector*, July 2007, pp. 1–15.
- Donato, Jeanne M. *Reducing Test Anxiety and Improving Academic Performance in Fourth Grade Students: Exploring an Intervention*. Southern Connecticut State University. ProQuest, <https://www.proquest.com/docview/501969636/abstract/9F798F6CC9454CC4PQ/1>. Accessed 9 Feb. 2022.
- Education, National Commission on Excellence in. "A Nation at Risk: The Imperative for Educational Reform." *The Elementary School Journal*, vol. 84, no. 2, 1983, pp. 113–30.
- English, David. *Proposed State Accountability Systems Under the Every Student Succeeds Act: A Summary of Fall 2017 Submissions*. Nov. 2017, pp. 1–5.
- Forte, Ellen. "Examining the Assumptions Underlying the NCLB Federal Accountability Policy on School Improvement." *Educational Psychologist*, vol. 45, no. 2, Apr. 2010, pp. 76–88. DOI.org (Crossref), <https://doi.org/10.1080/00461521003704738>.
- Fuhrman, Susan H., and Richard F. Elmore. *Redesigning Accountability Systems for Education*. Teachers College Press, 2002. DOI.org (Crossref), <http://choicereviews.org/review/10.5860/CHOICE.40-1678>.
- Fulton, Beth Ann. *The Relationship Between Test Anxiety and Standardized Test Scores*. Walden University. ProQuest, <https://www.proquest.com/docview/1786677753/abstract/6C77BEB970DC4074PQ/1>. Accessed 7 Feb. 2022.
- Gamson, David A., et al. "The Elementary and Secondary Education Act at Fifty: Aspirations, Effects, and Limitations." *RSF : Russell Sage Foundation Journal of the Social Sciences*, vol. 1, no. 3, 2015, pp. 1–29. baylor.primo.exlibrisgroup.com, <https://doi.org/10.7758/rsf.2015.1.3.01>.
- Hamilton, Laura S., et al. *Standards-Based Accountability Under No Child Left Behind: Experiences of Teachers and Administrators in Three States*. Rand Corporation, 2007.
- Haney, Walt. *The Myth of the Texas Miracle in Education*. no. 41, pp. 1–323.

- Hess, Frederick M., and Max Eden. *The Every Student Succeeds Act: What It Means for Schools, Systems, and States*. Harvard Education Press, 2017.
- Hochbein, Craig, et al. "Gamed by the System: Adequate Yearly Progress as an Indicator of Persistently Low-Achieving School Performance." *NASSP Bulletin*, vol. 97, no. 3, Sept. 2013, pp. 270–89. *SAGE Journals*, <https://doi.org/10.1177/0192636513479139>.
- Hunt, John W. "'A Nation at Risk' and No Child Left behind: Déjà Vu for Administrators?" *The Phi Delta Kappan*, vol. 89, no. 8, 2008, pp. 580–85.
- Jennings, Jennifer L., and Jonathan Marc Bearak. "'Teaching to the Test' in the NCLB Era: How Test Predictability Affects Our Understanding of Student Performance." *Educational Researcher*, vol. 43, no. 8, Nov. 2014, pp. 381–89. *SAGE Journals*, <https://doi.org/10.3102/0013189X14554449>.
- Jerald, Craig D. "The Hidden Costs of Curriculum Narrowing. Issue Brief." *Center for Comprehensive School Reform and Improvement*, Aug. 2006, pp. 1–6.
- Koretz, Daniel. "Adapting Educational Measurement to the Demands of Test-Based Accountability." *Measurement*, vol. 13, no. 1, Jan. 2015, pp. 1–25. *EBSCOhost*, <https://doi.org/10.1080/15366367.2015.1000712>.
- Koretz, Daniel M. *The Validity of Gains in Scores on the Kentucky Instructional Results Information System (KIRIS)*. RAND, 1998.
- Lam, Ricky. "Understanding Assessment as Learning in Writing Classrooms: The Case of Portfolio Assessment." *Iranian Journal of Language Teaching and Research*, 2018, pp. 19–36.
- Lowe, Patricia A., et al. "The Test Anxiety Inventory for Children and Adolescents (TAICA): Examination of the Psychometric Properties of a New Multidimensional Measure of Test Anxiety Among Elementary and Secondary School Students." *Journal of Psychoeducational Assessment*, vol. 26, no. 3, Sept. 2008, pp. 215–30. *SAGE Journals*, <https://doi.org/10.1177/0734282907303760>.
- Mary Lee Smith. *Put to the Test: The Effects of External Testing on Teachers*. 1991, pp. 8–11.
- McMurrer, Jennifer. "Choices, Changes, and Challenges: Curriculum and Instruction in the NCLB Era." *Center on Education Policy*, Dec. 2007, pp. 1–17.
- Moshe Zeidner. *Test Anxiety : The State of the Art*. Springer, 1998. *EBSCOhost*, <http://ezproxy.baylor.edu/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=nlebk&AN=69259&site=ehost-live&scope=site>.



- Mulvenon, Sean W., et al. *Impact of Accountability and School Testing on Students: Is There Evidence of Anxiety?* 2001, pp. 1–32. ERIC, <https://eric.ed.gov/?id=ED460155>.
- National Center for Education Statistics. *Changes in Instructional Hours in Four Subjects by Public School Teachers of Grades 1 Through 4*. US Department of Education.
- National Center on Intensive Intervention. “Handout 3: Examples of Evidence-Based Interventions.” *National Center on Intensive Intervention*, pp. 1–10.
- Newberg-Long, Donna. *Narrowing of Curriculum: Teaching in an Age of Accountability: A Phenomenological Study*. University of Denver, 2010. [baylor.primo.exlibrisgroup.com, https://search.proquest.com/docview/761133825?pq-origsite=primo](https://search.proquest.com/docview/761133825?pq-origsite=primo).
- Nussbaum, Martha Craven. *Not For Profit: Why Democracy Needs the Humanities*. New paperback edition., University Press, 2016.
- Phelps, Richard P. *Defending Standardized Testing*. L. Erlbaum Associates, 2005.
- Polikoff, Morgan S. “Instructional Alignment under No Child Left Behind.” *American Journal of Education*, vol. 118, no. 3, May 2012, pp. 341–68. DOI.org (Crossref), <https://doi.org/10.1086/664773>.
- Puma, Michael J., et al. *Exploring New Directions Title I in the Year 2000*. Distributed by ERIC Clearinghouse, 2000.
- Rhodes, Jesse H. *An Education in Politics: The Origins and Evolution of No Child Left Behind*. Cornell University Press, 2012.
- Riley, Richard W. “The Improving America’s Schools Act and Elementary and Secondary Education Reform.” *Journal of Law & Education*, vol. 24, no. 4, Fall 1995, pp. 513–66.
- Rothstein, Richard, and Rebecca Jacobsen. “The Goals of Education.” *Phi Delta Kappa International*, vol. 88, Dec. 2006, pp. 264–72.
- Sadovnik, Alan R., et al. *No Child Left Behind and the Reduction of the Achievement Gap: Sociological Perspectives on Federal Educational Policy*. Taylor & Francis Group, 2007. ProQuest Ebook Central, <http://ebookcentral.proquest.com/lib/bayloru/detail.action?docID=1474729>.
- Segool, Natasha K., et al. “Heightened Test Anxiety Among Young Children: Elementary School Students’ Anxious Responses to High-Stakes Testing.” *Psychology in the*

- Schools*, vol. 50, no. 5, 2013, pp. 489–99. *Wiley Online Library*, <https://doi.org/10.1002/pits.21689>.
- Shepard, Lorrie A. “What the Marketplace Has Brought Us: Item-by-Item Teaching With Little Instructional Insight.” *Peabody Journal of Education*, vol. 85, no. 2, Apr. 2010, pp. 246–57. *DOI.org (Crossref)*, <https://doi.org/10.1080/01619561003708445>.
- Triplett, Cheri Foster, and Mary Alice Barksdale. “Third through Sixth Graders’ Perceptions of High-Stakes Testing.” *Journal of Literacy Research*, vol. 37, no. 2, June 2005, pp. 237–60. *SAGE Journals*, [https://doi.org/10.1207/s15548430jlr3702\\_5](https://doi.org/10.1207/s15548430jlr3702_5).
- United States. “National Defense Education Act of 1958”, as Amended by the 88th Congress. Washington: U.S. Govt. Print. Off., 1964.
- United States Congress. *H.R.1 - 107th Congress (2001-2002): No Child Left Behind Act of 2001*. 8 Jan. 2002, <https://www.congress.gov/bill/107th-congress/house-bill/1>.
- United States Congress. *H.R.6 - 103rd Congress (1993-1994): Improving America’s Schools Act of 1994* | *Congress.Gov* | *Library of Congress*. 20 Oct. 1994, <https://www.congress.gov/bill/103rd-congress/house-bill/6>.
- United States Government Accountability Office. *Every Student Succeeds Act: Early Observations on State Changes to Accountability Systems*. United States Government Accountability Office, 2017.
- United States Congress. *[USC02] 20 USC 5812: National Education Goals*. Mar. 1994, <http://uscode.house.gov/view.xhtml?path=&req=granuleid%3AUSC-prelim-title20-section5812&f=&fq=&num=0&hl=false&edition=prelim>. Accessed 29 Aug. 2021.
- von der Embse, Nathaniel, et al. “Test Anxiety Interventions for Children and Adolescents: A Systematic Review of Treatment Studies from 2000-2010.” *Psychology in the Schools*, vol. 50, no. 1, Jan. 2013, pp. 57–71. *EBSCOhost*, <https://doi.org/10.1002/pits.21660>.
- Wang, Lihshing, et al. “Controversies of Standardized Assessment in School Accountability Reform: A Critical Synthesis of Multidisciplinary Research Evidence.” *Applied Measurement in Education*, vol. 19, no. 4, Oct. 2006, pp. 305–28. *EBSCOhost*, [https://doi.org/10.1207/s15324818ame1904\\_5](https://doi.org/10.1207/s15324818ame1904_5).
- Wisconsin Department of Public Instruction. *Sanctions for Title I Schools Not Making Adequate Yearly Progress (AYP)*. Apr. 2010.
- Zastrow, Claus von, and Helen Janc. *Academic Atrophy: The Condition of the Liberal Arts in America’s Public Schools*. Council of Basic Education, 2006.