

## **The Political Forces Driving High-Stakes Assessment in America's Public Schools**

Lindsey Caccavale, Marquita Hockaday, Leah Shy, and Chandra Stevens

In 2011, journalist Steven Brill published *Class Warfare: Inside the Fight to Fix America's Public Schools*. In this book, Brill narrated the story of educational reform in the United States. Following an assortment of key players in the arena, including a political advisor, hedge-fund managers, a software titan, a young college graduate, empowered parents, a documentary producer, and city mayors, Brill described what has led to the current high-stakes testing environment in America's public schools, the cornerstone of educational reform. Noticeably absent from the recount is the impact of educators, though stories of lawyer Randi Weingarten, president of the American Federation of Teachers, do appear. Overall, Brill's chronicle of recent shifts in policy documents educational change as coming from numerous political forces outside of the field of education, a report that is affirmed from the voices within education as well.

This educational change is most noticeably felt through the mandate of high-stakes testing, not simply due to the number and types of tests required, but more so due to the varied applications of their results. Popham (2004) stated "A high-stakes educational test is one that has serious consequences—either for the students who take the test or for those who prepared students for that test" (p. 63). Those consequences may include student promotion, tracking, and graduation; teacher salaries, retention, and firing; and school grading, closure, and takeover by the state (Popham, 2004). Indeed, where tests were once the province of educators, largely used to gauge student progress, they have now become the province of the government, used to gauge the efficacy of the entire public school system. A historical analysis of a number of government documents plots this confiscation.

Furthermore, a historical review of educational research and commentary demonstrates educators' objections to this shifting of assessments' territory from educators to policymakers since the pendulum first began to swing in that direction. In this article, we will use the historical perspective to examine the high-stakes testing phenomena and tell the story of how a nation's political machinery—in the name of student achievement, improved economy, and global competitiveness—wrested assessment from the hands of professional educators and attached high stakes to its applications. Next, we will highlight education professionals' responses to this seizure and the grounds on which they protested. Finally, we look at one state's recent conciliatory move in a slightly different direction and offer conjecture on what this move may signify.

### **Historical Perspective**

To best understand the phenomenon of high-stakes testing, it is crucial to understand the forces that have shaped the phenomenon as it currently exists. According to Edson (1986) "the context of historical inquiry, however, is not only the past—as many erroneously assumed—but the present as well" (p. 16). The historical researcher understands that meaning gathered about the present is gained through knowledge from the past. Also, Edson (1986) states, "historians both describe and analyze—they seek to understand as well as to explain the past" (24). Using historical perspective allows the researcher to analyze various artifacts, including primary sources, which may represent diverse perspectives in order to come to a formative conclusion about the present (Edson, 1986). Gall, Gall, and Borg (2007) define historical perspective as "a process of systematically searching for data to answer questions about a phenomenon from the past

to gain a better understanding of the foundation of present institutions, practices, trends, beliefs, and issues in education” (p. 529). Examining the current era of accountability and how it has been shaped by political and legislative mandates spanning the past thirty years provides a descriptive account of how high-stakes testing came to be the greatest measure of student achievement.

### **Historical Background**

The notion of testing students to gauge levels of achievement in the United States is a practice that has been in place since the colonial time period when students were given oral assessments, though it was not until the 19<sup>th</sup> century that students began to be evaluated on literacy skills due to a shift in curriculum and again in the 20<sup>th</sup> century when businesses began using scientific management models (Groen, 2012). Also, the work force began using tracking systems and education mimicked the business model, implementing measurements that focused on teacher effectiveness and school performance (Groen, 2012). This idea gained popularity as philosophers and education critics suggested tracking students’ test scores to rate the efficiency of teachers and force low performing schools to raise their standards (Barton, 1999; Groen, 2012). According to education critic Joseph Mayer Rice (1913) testing “would serve to raise the standard of the weaker schools to a plane of reasonable efficiency” (p. 277). The first surge of American public school systems relying on assessments that focused heavily on measurements and efficiency were in responses to the goals of “centralizing school administration, controlling school costs, and setting standards by which to compare school performance within and across districts” (Resnick, 1980, p. 3).

When President Lyndon B. Johnson gave his State of the Union Address in 1964, he was not focused on testing, but he did set a precedent for federal involvement in education. Johnson declared war on poverty in America and stated that our best weapon was our school system (Groen, 2012).

Therefore, in 1965 he developed the Elementary and Secondary Education Act (ESEA), the first national act that gave federal money to local school systems in an effort to provide for those students who were underserved in the past, such as minorities and students from a low socioeconomic status background (Popham, 2004). Prior to 1965, funding for public schools came only from local taxes (Popham 2004). However, with federal dollars came federal mandates.

ESEA included four goals: to make sure federal money was sent to local school districts that were impoverished; to meet the needs of academically disadvantaged students (at impoverished schools); to concentrate on creating state standards and tests that were aligned with said standards; and to make sure school systems were implementing measures to hold students accountable (Manna, 2011). However, ESEA was not without issues; there was no way to make sure local school districts were holding schools accountable for measuring test scores, the tests created were not effective, state test scores were not comparable, and Washington was not receiving clear information on whether or not federal aid was helping (Manna, 2011). Therefore, since the first implementation of ESEA, several presidents have modified the act because of its ambiguity in terms of accountability and measurement (Manna, 2011), thus intensifying the federal government’s involvement in education.

### **1981: Policymakers Outline the Risks of a Nation**

In 1981, Secretary of Education Terrell Bell, concerned with the widespread perception that American schools were performing poorly (Manna, 2011), organized the National Commission on Excellence in Education (NCEE) to study the quality of education in America. Bell’s worry was that underperforming schools implied the nation was somehow in jeopardy, especially when compared to other progressive countries (Manna, 2011). Thus, the Commission’s job was to provide concrete recommendations that could lead to

improvements in education, with its charter listing specific tasks, including evaluating teacher quality, comparing American schools and institutions of higher education with those of other advanced nations, and looking at how social and educational changes have impacted student achievement (NCEE, 1983). *A Nation at Risk*, the Commission's report, while not a federal mandate or federal legislation of any kind, would later influence education policy.

The members of the Commission recognized that the United States could no longer rely on older American ideals as the world became less isolated and more globalized. To determine if the education system in United States was in fact at risk, the Commission compared the achievement of American students with that of students from other advanced nations. Some of the data found by the Commission confirmed the underperformance of American students on academic tests when compared to other international industrialized nations (NCEE, 1983). More specifically, the Commission found that students' performances on aptitude assessments declined by 40 to 50 points from 1963 to 1980 (NCEE, 1983). While some of the data collected did revolve around the importance of various educational aspects (e.g., remediation, writing skills, and literacy), most of the emphasis was placed on defining achievement through test scores. Against the backdrop of the Cold War, this emphasis is of historical importance since this federal commission used assessment to criticize the state of the nation's public schools both by demarcating America's own educational decline and by comparing American schools to schools worldwide.

After the release of *A Nation at Risk*, state political leaders called for more intense education reforms and federal leaders began to focus policy decisions on matters of fairness and quality in education (Manna, 2011). According to Vinovskis (1998) the report's release also spurred political interest in the previously opposed development of assessment measures that would allow more accurate cross-state comparisons, specifically the

NAEP. In fact, at the 1984 Southern Regional Educational Board meeting, several governors including Arkansas governor Bill Clinton, advocated these measures (Vinovskis, 1998).

According to former Secretary of Education and educational historian Diane Ravitch (2010), an effort in the early 1990s to strengthen academic expectations gave way to conflicts of political ideology, and the fledgling standards movement made little headway, despite the fact that "the Clinton administration's Goal 2000 program gave the states federal money to write their own academic standards" (p. 19). The next big change would come with the next presidential administration, when "the standards movement was replaced by the accountability movement" (Ravitch, 2010, p. 16).

### **2001: The Federal Government Vows to Leave No Child Behind**

Perhaps no federal legislation created for educational purposes has been more controversial or more discussed than the No Child Left Behind (NCLB) Act ratified in 2001 under President George W. Bush. NCLB began as a bipartisan effort that created harmony across political groups for the sake of enforcing accountability and standardization in schools nationwide. The law was lengthy (at over 600 pages), including over 50 programs and the same four essential elements from ESEA (Manna, 2011). However, there are key differences between ESEA and NCLB. First, NCLB tried to correct the areas in which the legislators felt ESEA faltered: measurement and accountability (Manna, 2011). Second, ESEA and NCLB differ in the way in which they are (and in the case of ESEA *were*) enforced.

As a first step to correcting issues of accountability, federal policymakers under IASA required every state to create higher standards, nullify sporadic testing patterns and correct any tests that did not meet the new superior standards (Manna, 2011). NCLB policymakers continued the aforementioned practice, carrying it a step further

by implementing methods that enforced consequences on states that did not comply (Manna, 2011). Thus, local school divisions under the direction of each individual state implemented NCLB; however, local education agencies (LEAs) are at the mercy of the federal government when it comes to consequences for not meeting requirements outlined in NCLB. For instance, school choice is a common consequence that requires each LEA to give students the opportunity to attend a superior public school if the school they are currently attending is identified “for improvement, corrective action, or restructuring” (NCLB, 2002).

In order to determine student achievement, the issue of measurement also had to be rectified by the writers of NCLB. Thus, according to the parameters set in NCLB, a states’ adequate yearly progress (AYP) became the tool used to measure student success. Among other things, AYP is based on a state’s assessment results. School districts, and schools, are either given rewards such as recognition and monetary bonuses or consequences for making or exceeding their AYP objectives and goals, or failing to make AYP over a number of years (NCLB, 2002). Therefore, the framers of NCLB defined success by students’ assessment results on a grander scale than any other federal legislation in history.

### **Under the Current Administration**

Ravitch noted that many anticipated positive change once the Bush Administration was out of office. She wrote on her blog: “I thought President Obama would ditch high-stakes testing and federal sanctions and chart a new course. He didn’t. He built on the foundation of NCLB and made the stakes even higher by tying teacher evaluations to test scores” (Ravitch, 2013a, para. 3-4).

Obama’s Secretary of Education Arne Duncan built on the foundation of NCLB by creating Race to the Top (RTTT), which according to the U.S. Department of Education is:

a competitive grant program designed to encourage and reward States that are creating the conditions for education innovation and reform; achieving significant improvement in student outcomes, including making substantial gains in student achievement, closing achievement gaps, improving high school graduation rates, and ensuring student preparation for success in college and careers; and implementing ambitious plans in four core education reform areas. (2009, p. 1)

RTTT utilized federal stimulus money from the American Recovery and Reinvestment Act of 2009 in order to entice states to compete. In addition, it offered states the opportunity to apply for waivers from NCLB’s requirements, provided they complied with the administration’s policy agenda for reform, one that relies heavily on high-stakes testing as its accountability backbone.

In December of 2013, Duncan delivered a speech entitled, *The Threat of Educational Stagnation and Complacency*. In this speech, spurred by the release of the 2012 Program for International Student Assessment (PISA), Duncan (2013) utilized some of the same rhetoric first heard in *A Nation at Risk*, calling the U.S. performance on PISA “a picture of educational stagnation” (para. 9) and continued: “That brutal truth, that urgent reality, must serve as a wake-up call against educational complacency and low expectations” (para. 10). He notes that while graduation rates are rising, America’s 15- and 17-year-old students are not making progress when compared to peers in other high-performing countries, like Shanghai, Singapore, and Korea. He also debunks the notion that our underperforming gap groups are the cause for America’s ranking, stating that our 15-year-old white students are not among the world’s top performers. Ultimately, Duncan recommends that we learn from other countries to improve our education model, while pinpointing “adoption of the Common Core, and the Race to the Top program, as two critical steps in making the U.S. more like high-performing nations” (para. 45). Consequently,

Duncan and other stakeholders shaping education policy continue to make a case for more high-stakes accountability measures.

### **The Perspective of Educators**

Thus far, we have chronicled the 30-year political evolution of standards and accountability monitored via high-stakes standardized tests. It is important to note that this evolution was never without censure. Indeed for decades, education professionals have urged caution in the development and application of assessments.

The National Center for Fair and Open Testing (FairTest) has advocated the use of authentic assessments over standardized assessment since 1985, just two years after the release of *A Nation at Risk*. In 1989, the executive director of FairTest Dr. Monty Neill and colleague Noe J. Medina (1989) reported that our nation's overuse and misuse of standardized tests narrows what is taught, limits student opportunity, and deprives the public of its power to determine what matters. Citing the presence of bias and the lack of construct validity and reliability inherent in all tests, the authors contended that standardized tests often do greatest damage to minority and poor students (Neill & Medina, 1989). Ironically, NCLB was founded on the premise that requiring such measures would improve schooling for these subgroups. Rather than improving schooling, however, high-stakes standardized testing reduces that which is taught to that which is tested: easily measured constructs that do not constitute the richness of literacy in math nor in language arts (Neill & Medina, 1989). Indeed, as Neill (2006) pointed out, "the higher the stakes, the more schools focus instruction on the tests" (p. 30).

In more recent years, University of Oregon Professor Yong Zhao has also critiqued high-stakes testing, tied to the standardization movement, as a dangerous extension of the federal government into matters it is not prepared to effectively oversee, and which it uses for its own agenda. In 2009, he critiqued NCLB in how it narrowed the definition of

student success: "While the intention is to ensure every child receives a good education, the problem is that NCLB practically defines good education as being able to show good scores in a limited number of subjects" (p. x).

Moreover, if ESEA marked the government's first attempt to utilize federal mandates and federal money to steer state and local school policy, RTTT marked its best attempt. Al Ramirez (2010), professor at University of Colorado-Colorado Springs, clarified how the high-stakes testing requirement in RTTT has led to more federal control over states' policies and practices, stating "about 10 percent [of funding] comes from the federal government. Despite this lopsided funding arrangement, federal policy now reaches into every school district and classroom in the country" (para. 3). This situation is problematic since national government involvement in state and local school policy destabilizes these systems (Ramirez, 2010).

Likewise, Ravitch expressed concern about the motives behind the decisions driving education further into a high-stakes testing environment. "President Obama's Race to the Top adopted the same test-based accountability as NCLB. The two programs differed in one important respect: where NCLB held schools accountable for low scores, Race to the Top held both schools and teachers accountable." (Ravitch, 2013b, p. 99) She continued to describe the problems with holding individual teachers accountable to single-year test scores and value-added measures, as Race to the Top requires states to do. "Stated as politely as possible, value-added assessment is bad science. It may even be junk science. It is inaccurate, unstable, and unreliable" (Ravitch, 2013b, p. 113). Research evidence being conducted by researchers and scholars in the field does not support the high-stakes testing policies.

Ironically, within the classroom, teachers respond to what Invernizzi, Landrum, Howell, and Warley (2005) referred to as "unprecedented political insistence on the use of research-based, scientifically proven assessments and instructional

techniques” (p. 610) even when these measures fall short in providing them with what they need to improve daily instruction. To combat these shortcomings, teachers continue to use instructionally informative assessments of their own, merely adding these to the mandated high-stakes assessments (Invernizzi, Landrum, Howell, & Warley, 2005). The trade-off of more assessments, however, is less instructional time.

### **Politicians and Educators Begin to Work Together in Virginia**

Recognizing instruction as paramount to student learning and achievement, in 2010, Virginia Governor Bob McDonnell requested the Joint Legislative Audit and Review Commission (JLARC) conduct a study on ways to promote and ensure early reading proficiency and comprehension among third graders in Virginia public schools (Tittermary, 2011). In 2011, the JLARC staff released their report entitled, *Strategies to Promote Third Grade Reading Performance in Virginia*. In order to help schools focus on the development of reading skills in third grade, the JLARC study recommended that the Standards of Learning (SOL) tests taken by third grade students be limited to reading and math only. As one of only four states that did not originally adopt the Common Core State Standards as state curriculum from the outset (Baker, 2014), Virginia has slightly more freedom over educational initiatives than most other states. Virginia Superintendent of Public Instruction, Dr. Steve Staples, asked teachers to “collaborate and take full advantage of this opportunity to design assessments that support instruction and assess 21<sup>st</sup> century skills as well as SOL content” (VDOE, 2014, para. 6).

In 2014, Virginia approved new legislative mandates outlined in *Guidelines for Local and Alternative Assessments for 2014-2015*, amending §22.1-253.13:3.C of the *Code of Virginia* to eliminate several Standards of Learning (SOL) tests (VDOE, 2014). New *Virginia Code* specifically states:

The Standards of Learning assessments administered to students in grades three through eight shall not exceed (a) reading and mathematics in grades three and four; (b) reading, mathematics, and science in grade five; (c) reading and mathematics in grades six and seven; (d) reading, writing, mathematics, and science in grade eight; and (e) Virginia Studies and Civics and Economics once each at the grade levels deemed appropriate by each local school board (para. 2).

Although the Virginia Department of Education (VDOE) decreased the number of subjects assessed by high-stakes end-of-course SOL assessments, the new legislation requires that each local school system develop alternate assessments for each subject area in which the SOL assessment was eliminated. The VDOE (2014) states that the new guidelines intend to eliminate some of the high-stakes assessments used for accountability and instead, encourage greater use of assessments designed to inform instruction.

### **Conclusion**

In sum, an analysis of the political forces, policies, and studies over the previous three decades clarifies the driving forces behind the current high-stakes testing environment in America’s public education system. While at a national level we appear to be repeating history, policy makers in Virginia are hedging down a different path. As Invernizzi, Landrum, Howell, and Warley (2005) suggested, perhaps a “peaceful coexistence of test developers, policy makers, and teachers in an era of accountability” (p.610) can be established. Carol Ann Tomlinson (2008), master teacher and University of Virginia professor, wrote in a personal memoir:

Lorna Earl (2003) distinguishes between assessment of learning, assessment for learning, and assessment as learning. In many ways, my growth as a teacher slowly and imperfectly followed that progression. I began seeing assessment as judging

performance, then as informing teaching, and finally as informing learning. In reality, all those perspectives play a role in effective teaching. The key is where we put the emphasis. (p. 13)

Perhaps America is on a journey similar to the one Tomlinson describes in her memoir. While the majority of the country seeks high-stakes assessment as a means for judging student and teacher performance, Virginia's latest policy supports the importance of instructional time and formative assessment. Perhaps as Virginia educators, policy makers, and researchers continue to work together and set the stage for new research in the field of education pertaining to assessment, the Nation will follow suit.

#### References

- Baker, C. R. (2014, January 4). Opting out of the common core. *Desert News*. Retrieved from <http://www.deseretnews.com/article/865593461/Opting-out-of-the-Common-Core.html?pg=all>
- Barton, P. (1999, March). *Too much testing of the wrong kind; too little of the right kind in K-12 education: A policy information perspective*. Princeton, NJ: Policy Information Center.
- Brill, S. (2011). *Class warfare: Inside the fight to fix America's schools*. New York, NY: Simon & Schuster.
- Duncan, A. (2013, December 3). The threat of educational stagnation and complacency. Retrieved from <http://www.ed.gov/news/speeches/threat-educational-stagnation-and-complacency>.
- Edson, C. H. (1986). Our past and present: Historical inquiry in education. *Journal of Thought* 21(3), 13-27.
- Gall, M. D., Gall, J. P., & Borg, W. R. (2007). *Educational Research: An Introduction*. (8<sup>th</sup> Ed.). Boston, MA: Pearson Education, Inc.
- Groen. M. (2012). NCLB-the educational accountability paradigm in historical perspective. *American Educational History Journal*, 39(1), 1-14.
- Hunt, J.W. (2008). A nation at risk and No child left behind: Déjà vu for administrators? *The Phi Delta Kappan*. 89(8), 580-585.
- Invernizzi, M. A., Landrum T. J., Howell, J. L., & Warley, H. P. (2005). Toward the peaceful coexistence of test developers, policymakers, and teachers in an era of accountability. *The Reading Teacher*. 58(7), 610-618. doi:10.1598/RT.58.7.2
- Joint Legislative Audit and Review Commission. (2011, November). Report of Strategies to Promote Third Grade Reading Performance in Virginia. (State Document No.7). Retrieved from <http://jlarc.virginia.gov/reports.shtml>
- Manna, P. (2011). *Collision course: Federal education policy meets state and local realities*. Washington DC: CQ Press.
- Neill, M. (2006). The case against high-stakes testing. *Principal*, 85(4), 28-32.
- Neill, M., and Medina, N. J. (1989). Standardized testing: Harmful to educational health. *The Phi Delta Kappan*, 70(9), 688-697.
- Popham, W. J. (2004). *America's "failing" schools: How parents and teachers can cope with No Child Left Behind*. New York, NY: RoutledgeFalmer.
- Ramirez, A. (2012, April 26). It's time to reform federal education policy. *Education Week*.

- Retrieved from <http://www.edweek.org/ew/articles/2010/04/28/30ramirez.h29.html?tkn=OXMF3kgFuq4edM90cNFbLIUghYlaxaKnPPA3&print=1>
- Ravitch, D. (2010). *The death and life of the great American school system: How testing and choice are undermining education*. New York, NY: Basic Books.
- Ravitch, D. (2012, September 1). About that meeting at the White House. [Blog post]. Retrieved from <http://dianeravitch.net/2012/09/01/about-that-meeting-at-the-white-house/>
- Ravitch, D. (2013a, August 20). Should the U.S. government monitor people like me? [Blog post]. Retrieved from <http://dianeravitch.net/2013/08/20/should-the-u-s-government-monitor-people-like-me/>
- Ravitch, D. (2013b). *Reign of error: The hoax of the privatization movement and the danger to America's public schools*. New York, NY: Random House.
- Resnick, D. (1980). Minimum competency testing historically considered. *Review of Research in Education* (pp. 3-29). Retrieved from: <http://www.jstor.org/stable/1167122>
- Rice, J. M. (1913). *Scientific management in education*. New York, NY: Hinds, Noble and Eldredge.
- Rothstein, R. (2011, September 26). A bet over No Child Left Behind. [Blog post]. Retrieved from <http://www.epi.org/blog/rothstein-ravitch-no-child-left-behind/>
- The National Commission on Excellence in Education. (1983). *A nation at risk: The imperative for educational reform*. Washington DC: The Commission.
- The No Child Left Behind Act of 2001, 147 U.S.C. §§107-110 (2002).
- Tittermary, G. S. (2011, November 4). [Letter to Honorable Charles J. Colgan]. General Assembly Building, Richmond, VA.
- U.S. Department of Education. (2009, November). *Race to the Top executive summary*. Retrieved from <http://www2.ed.gov/programs/racetothetop/executive-summary.pdf>
- Vinovskis, M. A. (1995). *Education, society, and economic opportunity: A historical perspective on persistent issues*. New Haven, CT: Yale University Press.
- Vinovskis, M. A. (1998). *Overseeing the nation's report card: The creation and evolution of the National Assessment Governing Board (NAGB)*. Retrieved from <http://www.nagb.org/content/nagb/assets/documents/publications/95222.pdf>
- Virginia Board of Education. (2014, September 18). Guidelines for local alternative assessments. *Virginia Board of Education*. Retrieved from [http://www.doe.virginia.gov/boe/meetings/2014/07\\_jul/agenda\\_items/draft\\_local\\_assessment\\_guidelines.pdf](http://www.doe.virginia.gov/boe/meetings/2014/07_jul/agenda_items/draft_local_assessment_guidelines.pdf)
- Virginia Department of Education. (2014, September 18). Board of Education approves local test guidelines. *Virginia Department of Education News*.
- Zhao, Y. (2009). *Catching up or leading the way: American education in the age of globalization*. Alexandria, VA: Association



for Supervision and Curriculum Development.

Zhao, Y. (2010, May 23). Mass localism: How might the Race to the Top money be better spent? [Blog post]. Retrieved from <http://zhaolearning.com/2010/05/23/mass-localism-how-might-the-race-to-the-top-money-be-better-spent/>