

The Unseen Obstacles: Union and Testing Company Reactions to Honest Testing Initiatives Under Clinton and Bush

Executive Summary

This report examines the enduring challenges to educational accountability in the United States, focusing on the reactions of teachers' unions and testing companies to proposals for "honest testing" during the presidencies of Bill Clinton and George W. Bush. "Honest testing" is defined herein as assessments characterized by strict security protocols, regularly updated national norms, consistent rotation of test questions, and administration by independent, external proctors. These measures are critical to prevent score inflation and the detrimental practice of "teaching to the test." This stands in stark contrast to the pervasive "Lake Wobegon Effect," a phenomenon where nearly all states and school districts report "above average" student performance, primarily due to lax test security, reliance on outdated norm groups, and various deceptive reporting methods. The continued relevance of these issues, as indicated by a 2025 press release calling to "End U.S. Testing Chaos Denying Parents Honest Scores" , underscores the deeply entrenched nature of the problem.

The analysis reveals that resistance to genuine testing reform manifested through several key avenues during the Clinton and Bush eras. These included active political opposition to national test proposals by legislative bodies, strategic institutional inaction and policy manipulation by state agencies, and the perpetuation of profitable, yet flawed, testing systems by commercial entities and their academic consultants. These forms of resistance collectively contributed to the persistence of inflated scores and the undermining of true educational accountability, highlighting a complex interplay of economic, political, and professional self-interest that has consistently impeded meaningful reform.

Introduction: The "Lake Wobegon" Challenge to Educational Accountability

Defining the "Lake Wobegon Effect": Inflated Scores, Outdated Norms, and Deceptive Practices

The "Lake Wobegon Effect" describes a statistically improbable phenomenon in American education where virtually all states and school districts claim to be "above the national average" on standardized achievement tests. By early 1988, all 50 U.S. states were reporting scores above the publisher's national norm. A subsequent 1989 survey confirmed this widespread anomaly, finding that 48 of the 50 states, 90% of elementary schools, and 80% of secondary schools continued to exceed the national norm. This paradoxical situation arises because scores are typically compared against an old "norm group" from the past, rather than a current

national average of all test-takers.

John Jacob Cannell's seminal 1987 and 1989 reports meticulously documented how this effect was sustained through a combination of outright cheating by educators, deceptive testing practices, and misleading reporting methods. Instances of cheating included teachers changing students' answer sheets after tests, providing more than the allotted time, using exact test questions for review, or making copies of tests for students. Administrators were also implicated, sometimes forcing teachers to teach items known to be on the test, linking high scores to promotions. Deceptive practices extended to excluding low-performing students, using test preparation materials that were virtually identical to actual test questions, and employing "test-curriculum alignment" strategies that narrowed the curriculum to match test content. Misleading reporting methods further obscured true achievement, such as manipulating statistical presentations like stanines or percentile ranks to make scores appear higher. The pervasive nature of the "Lake Wobegon Effect" across nearly all states and districts indicated a deeply systemic issue rather than merely isolated instances of misconduct. This widespread score inflation was not simply a statistical anomaly; it was a product of an educational ecosystem that incentivized and enabled deceptive practices. This systemic benefit is crucial for understanding the strong resistance encountered by proposals for "honest testing," as such reforms inherently threatened the established order and would expose actual performance levels.

The Call for "Honest Testing": Characteristics of Valid, Secure, and Accurate Assessments

In response to the "Lake Wobegon Effect," advocates like Cannell championed "honest testing," characterized by stringent security measures designed to prevent manipulation and ensure accurate assessment. These measures include individually sealed test booklets, delivery of tests to teachers only on the day of administration, the mandatory use of independent, outside test proctors, and the yearly rotation of a substantial portion of test questions to prevent familiarity and "teaching to the test". Unlike "Lake Wobegon" tests, which allow for nearly all students to appear above average, legitimate standardized tests should, by design, reflect a normal distribution where only about 50% of students score above the current national average. Furthermore, honest tests should utilize current national norms, ensure the inclusion of all students (including special education and bilingual learners) in reported scores, and employ technological methods like scanning for suspicious erasures or cluster variance to detect irregularities. Cannell also advocated for holding test publishers legally liable for selling deceptive tests or for practices that undermine test validity. The National Assessment of Educational Progress (NAEP) was often cited as an example of a legitimate standardized test due to its robust security and external oversight.

The Political and Educational Landscape at the Onset of Testing Reform Efforts

The national educational landscape in the 1980s was significantly shaped by reports such as "A Nation at Risk" (1983), which highlighted declining educational achievement and fueled public demand for greater accountability in schools. Paradoxically, while national reports painted a pessimistic picture, local "standardized" test scores continued to present an optimistic and contradictory message of improvement.

Tests originally intended as mere "instructional aids" rapidly transformed into "accountability yardsticks" due to mounting pressure from the public, school boards, and state legislatures. This fundamental shift in purpose, however, was unaccompanied by necessary reforms in test design and security. The result was the creation of a "high stakes" environment where educators faced immense pressure to produce high scores, often leading to widespread manipulation and the perpetuation of the "Lake Wobegon Effect".

III. The Clinton Administration and the Pursuit of National Standards

President Clinton's Engagement with Testing Integrity Concerns, Spurred by Findings like Cannell's

In April 1990, while serving as Governor of Arkansas, Bill Clinton was confronted by John Jacob Cannell regarding the unbelievably high and politically flattering test scores in Arkansas, a state known for its high illiteracy rates. Governor Clinton initially denied any knowledge of widespread cheating within the state's education system.

However, following a front-page newspaper exposé of Cannell's charges, Governor Clinton personally contacted Cannell. During their conversation, Clinton actively sought advice on how to "preserve the integrity of the test" in Arkansas. Cannell provided key recommendations for achieving genuine test validity: changing test questions annually, maintaining a large bank of questions, promoting a broad curriculum, testing infrequently, and discouraging excessive test preparation. Subsequently, Arkansas announced plans for improvements in test security, indicating a direct response to the concerns raised.

Clinton's 1996 Proposal for a National Achievement Test with Strict Security

As President, Bill Clinton later championed the idea of a national achievement test with strict security in 1996. This proposal aimed to introduce a more reliable and comparable measure of student achievement across the nation, aligning with the principles of "honest testing" that Cannell had advocated.

Instance 1: Congressional Opposition to Clinton's National Test Proposal

President Clinton's 1996 proposal for a national achievement test, intended to introduce greater integrity and comparability to educational assessment, was explicitly "refused by the Republican Congress". This refusal constituted a direct political barrier to implementing a national, secure testing system.

Opponents articulated various concerns for their rejection. They argued that U.S. students were already subjected to excessive testing, likening the proposed additional test to "fatten[ing] cattle by weighing them more often". They also warned that a national test would inevitably lead to "teaching to the test," resulting in a "watered down" curriculum and "drill-and-kill" instructional methods, which they believed would ultimately reduce genuine student motivation. A prominent political objection centered on the fear that such a test would establish a "de facto national

curriculum through the back door" and thereby undermine the traditional authority of state and local school officials over education.

The congressional rejection of Clinton's national test, while publicly justified by pedagogical and local control arguments, effectively maintained the decentralized and easily manipulable state-level testing landscape. By blocking a federally mandated, secure assessment, this political resistance inadvertently perpetuated the conditions that enabled the "Lake Wobegon" effect, thereby preventing a significant step towards "honest testing" at a national scale. This illustrates how resistance can be cloaked in seemingly legitimate concerns to protect a beneficial (for some) status quo. The arguments against Clinton's proposal, regardless of their theoretical validity, functioned to prevent a reform that would have exposed the true, lower performance levels enabled by the existing, less secure state tests.

Testing Companies' Systemic Resistance (Early Era)

Commercial test publishers actively resisted measures that would ensure test integrity and transparency during this period. They explicitly "refused" to provide Cannell with the necessary data for a proper statistical analysis of their test scores, hindering independent oversight. They also denied knowledge of the total number of children tested nationwide and individual state and district results, further contributing to data opacity.

The resistance extended to aggressive tactics against critics. McGraw-Hill, a prominent test publisher, "threatened legal action" against Cannell for daring to publish his findings, demonstrating an attempt to silence critics and protect their business interests. This move went beyond passive non-cooperation to active intimidation, highlighting the high stakes involved for testing companies in maintaining the profitable "Lake Wobegon" status quo.

Publishers consistently perpetuated deceptive norms by selling "older" tests and continuing to use outdated "norm groups" for comparison. This practice inherently inflated current student scores, making it statistically possible for nearly all students to appear "above average". They further facilitated score inflation by offering specialized "low socioeconomic norms" and "large city norms" to districts, allowing even struggling schools to claim "above average" status.

A critical aspect of this systemic resistance was the financial relationship between publishers and academics. Publishers paid "consultant fees" to university and public educators, whom Cannell described as "mere sycophants" who provided academic legitimacy to these commercially driven, often deceptive, practices. Furthermore, it was later exposed that some academics consulting for publishers also operated "side businesses" that provided "crib-sheet-like test preparation materials" containing actual test answers, directly undermining the validity of the tests they helped develop, for profit. This financial incentive created a profound conflict of interest, as these academics were positioned to benefit from the perpetuation of flawed testing practices.

The multi-faceted resistance from testing companies, encompassing legal threats, data opacity, and the co-optation of academic expertise, illustrates a powerful defense of a lucrative business model. This active and strategic opposition to transparency and genuine test integrity ensured the perpetuation of inflated scores, which were beneficial to their clients (schools and administrators) and, by extension, to their own profits. This deeply embedded commercial interest served as a significant barrier to the adoption of "honest testing" practices, as such changes would directly threaten their market and influence.

IV. The Bush Administration: No Child Left Behind and

Persistent Integrity Issues

President George W. Bush's Focus on Accountability Through the No Child Left Behind Act (NCLB)

Building on his emphasis on standards and testing as Governor of Texas, George W. Bush enacted the No Child Left Behind Act (NCLB) in 2002. This landmark legislation significantly expanded the federal government's role in education, mandating annual standardized testing and accountability measures for schools.

However, a critical policy choice under Bush was his opposition to a *national* test designed in Washington D.C. Instead, he preferred that states develop and administer their own yearly tests to preserve local control. This decision, while framed as supporting local autonomy, left the existing, vulnerable state-level testing infrastructure largely intact and susceptible to the very manipulations Cannell had exposed. Bush's stance of opposing a national test while mandating state yearly testing under NCLB created a policy vacuum that allowed the "Lake Wobegon" effect to persist and even worsen. By decentralizing test design and implementation, his administration implicitly relied on the same flawed state-level infrastructure that Cannell had exposed, effectively resisting the implementation of truly "honest" (i.e., secure, externally monitored, nationally normed) testing.

The "Texas Miracle" and its Subsequent Scrutiny for Inflated Gains

The "Texas Miracle"—a narrative of dramatic improvements in Texas school achievement that was central to Bush's political ascent—came under significant scrutiny. Independent analyses by the Rand Corporation and Professor Walter Haney concluded that most of these reported gains were "bogus" and "more illusory than real" when compared to results from the National Assessment of Educational Progress (NAEP). Further investigations, such as by the Dallas Morning News, uncovered widespread cheating in over 200 Texas schools, highlighting the extent of manipulation within the state's testing system.

Instance 2: Texas Education Agency's Passive Approach to Cheating Detection

During George W. Bush's governorship and extending into his presidency, the Texas Education Agency (TEA) demonstrated a notable lack of proactive effort in detecting cheating. Despite the availability of basic statistical techniques, such as cluster analysis to identify suspicious response patterns or optical scanning for suspicious erasures, the TEA "chose not to use it" proactively. Furthermore, while the agency did perform erasure analysis, it failed to act on this crucial information "unless they get a complaint!". This institutional inaction effectively allowed widespread dishonest testing practices to persist unchecked.

The TEA's deliberate passivity in utilizing readily available cheating detection methods represents a clear form of institutional resistance to "honest testing." By choosing not to proactively investigate, the agency implicitly facilitated the perpetuation of inflated test scores. This inaction served to bolster the politically favorable "Texas Miracle" narrative, prioritizing the appearance of success over genuine accountability and the exposure of educational deficiencies. The choice to avoid uncovering cheating, despite the availability of tools,

demonstrates a deliberate decision to allow dishonest testing to continue, thereby maintaining the illusion of progress.

Instance 3: Manipulation of Student Exclusion Policies to Inflate Scores

During Bush's tenure, particularly around 2004, the Texas Education Agency (TEA) implemented a policy that led school districts to deny special education services to thousands of students. This policy was influenced by pressure from the Bush administration to reduce the number of special education students, as their test scores were often excluded from school ratings, and high exclusion rates were perceived as "gaming the system" under NCLB. While federal law allowed only 1-3% exclusion for severe disabilities, Texas excluded 9% of its students from accountability metrics. This practice, famously dubbed the "ugly Texas two-step" by NCLB architect Sandy Kress, artificially inflated school performance metrics by removing low-scoring students from the pool of those whose scores counted towards official ratings. This manipulation of student exclusion policies represents a highly sophisticated and ethically problematic form of systemic resistance to honest testing. It goes beyond direct cheating on tests to manipulating the very population being assessed, thereby distorting the true picture of educational achievement. This reveals a deep institutional incentive to achieve favorable accountability metrics under high-stakes policies like NCLB, even at the expense of vulnerable students and the integrity of the data. The active removal of low-performing students from accountability calculations directly boosted reported scores, serving the clear motivation to meet NCLB's high-stakes demands and maintain a positive political narrative.

Teachers' Unions' Evolving Stance on High-Stakes Testing under NCLB

The National Education Association (NEA) and the American Federation of Teachers (AFT) initially adopted a stance of neutrality regarding the No Child Left Behind Act. However, their positions evolved into significant criticism as NCLB's implementation progressed. The NEA, in particular, engaged in "public denunciation" of the law, while the AFT pursued a "more considered" approach.

The unions' criticisms primarily focused on the practical consequences of NCLB's high-stakes environment within what Cannell identified as a corrupt testing infrastructure, rather than direct opposition to test integrity itself. Their concerns included the inadequacy of federal resources to meet the law's mandates, the reduction of valuable instructional time due to excessive test preparation ("teaching to the test"), and the inaccurate identification and measurement of English language learners and students with disabilities. They also voiced frustration over what they perceived as federal infringement on state and local control over education.

While teachers' unions did not fundamentally oppose the principle of honest testing (as evidenced by their initial positive reactions to Cannell's findings on "rigged scales"), their resistance to NCLB's implementation indirectly contributed to the persistence of the "Lake Wobegon" effect. By focusing their advocacy on mitigating the negative impacts of high-stakes testing on educators (e.g., teaching to the test, resource constraints) rather than demanding fundamental test integrity reforms (e.g., external proctors, question rotation), their efforts, while protecting their members, did not fully address the underlying systemic corruption that allowed scores to be inflated. This meant that the "dishonest" scores continued, creating a form of

indirect resistance to true honest testing by focusing on symptoms rather than root causes.

V. Analysis of Resistance: Motivations and Consequences

The Economic Incentives Driving Testing Companies' Resistance to Genuine Reform

The standardized testing industry operates as a multi-million dollar business. Companies derive significant profits from selling tests that consistently produce flattering "above average" scores, as this outcome encourages continued adoption by school districts and administrators who seek positive public relations and evidence of success. The practice of using unchanging test questions and outdated norms guarantees a steady upward trend in scores, which translates into "good news" for schools and administrators, thereby reinforcing the market demand for these particular tests.

Resistance to fundamental reforms is primarily driven by a desire to protect these lucrative business practices. This defense can involve aggressive tactics, such as threatening legal action against critics who expose flaws, or strategically co-opting academic experts to lend credibility to their flawed products. The financial incentive to maintain a system that produces flattering, albeit misleading, results is a powerful driver of resistance to genuine test integrity.

Political Motivations Behind Opposition to National Testing and Robust Accountability

Political leaders, from governors to presidents, benefit significantly from inflated test scores, as these allow them to claim "educational miracles" and demonstrate effective leadership and policy success. This creates a strong incentive to maintain a testing system that can produce such politically favorable outcomes.

Opposition to national tests, exemplified by the Republican Congress's refusal of Clinton's proposal and President Bush's preference for state-level tests, was often framed in terms of preserving local control and preventing federal overreach. However, these arguments also served to protect a decentralized testing environment where local scores could be more easily managed and manipulated for political advantage. The existence of a "conspiracy of silence," as described by a former Virginia state testing director, highlights a deliberate effort by administrators and politicians to suppress revelations of deficiencies and maintain a positive public perception of educational success.

The Complex Position of Teachers' Unions: Advocating for Educators While Navigating Accountability Demands

Initially, teachers' unions, through their leadership, publicly acknowledged the inherent illogic and "rigged scales" of the "Lake Wobegon" effect. This indicated that, in principle, they were not opposed to the idea of honest testing but rather to the lack of integrity in the existing system. However, under the intensified "high stakes" environment of NCLB, their primary focus shifted towards protecting educators from undue pressure, curriculum narrowing, and inadequate resources. This led to criticisms of the *implementation* of accountability policies that, without

genuine test integrity, placed unfair burdens on teachers and distorted instructional practices. While advocating for their members, this stance often resulted in an indirect resistance to the fundamental test integrity reforms (such as external proctors or question rotation) that might have exposed lower true scores and potentially increased pressure on educators. Their focus on the symptoms of high-stakes testing, rather than the root causes of test dishonesty, implicitly allowed the flawed testing infrastructure to persist.

The collective self-interest of these diverse stakeholders—testing companies prioritizing profit, school administrators seeking positive public image and career advancement, politicians aiming for favorable narratives, and teachers' unions striving to protect their members from perceived unfairness—created a powerful, multi-layered, and self-reinforcing system of resistance to "honest testing." This complex interplay of motivations explains why fundamental reforms have been so difficult to implement and why the "Lake Wobegon" effect has persisted for decades. Each group has a reason to prefer the current, manipulable system or to resist changes that expose true performance. This creates a powerful, systemic inertia, where the resistance is not monolithic but a dynamic interplay of economic, political, and professional self-preservation, which collectively acts to impede genuine testing reform.

VI. Conclusion: Lessons Learned and Future Directions for Honest Testing

Recap of the Persistent Challenges and the Three Identified Instances of Resistance

The "Lake Wobegon Effect," characterized by artificially inflated scores and deceptive testing practices, has remained a deeply entrenched and persistent challenge in American education. Its continued relevance, as highlighted by contemporary calls for "honest scores" in 2025, underscores the enduring nature of the problem, decades after its initial exposure by John Jacob Cannell.

During the presidencies of Bill Clinton and George W. Bush, specific instances of resistance to "honest testing" initiatives were clearly identifiable:

1. **Congressional Refusal of President Clinton's National Achievement Test Proposal (1996):** This significant political opposition, primarily from the Republican Congress, cited concerns over local control, curriculum influence, and perceived over-testing. This resistance effectively blocked a national move towards a secure, comparable testing system, thereby preserving the decentralized and manipulable state-level testing landscape.
2. **Texas Education Agency's Passive Approach to Cheating Detection and Active Manipulation of Student Exclusion Policies (Bush Era):** Under Governor and later President Bush, the Texas Education Agency demonstrated institutional resistance through inaction. This included failing to proactively use available cheating detection methods and actively manipulating student exclusion policies, particularly for special education students, to artificially inflate reported scores and meet accountability targets under NCLB. This deliberate avoidance of uncovering cheating and the manipulation of student populations directly undermined the integrity of assessment data.
3. **Testing Companies' Continued Perpetuation of Flawed Practices and Co-optation of Academics:** Throughout both administrations, testing companies systematically resisted

genuine reform. This resistance manifested through their continued use of outdated norms, refusal to provide data transparency to researchers, issuing legal threats against critics like Cannell, and benefiting from academics who developed and promoted "crib-sheet-like" test preparation materials that directly undermined test validity for profit. This systemic resistance prioritized economic gain over assessment integrity.

Table 2: Key Instances of Resistance to Honest Testing (Clinton & Bush Administrations)

Administration	Initiative/Policy for Honest Testing	Stakeholder(s) Resisting	Specific Action of Resistance
Clinton (1993-2001)	National Achievement Test with Strict Security (1996 proposal)	Republican Congress	Refusal to pass legislation, citing concerns over federal overreach, curriculum narrowing, and over-testing
Bush (2001-2009)	NCLB's accountability framework (reliant on state tests)	Texas Education Agency (TEA)	Passive approach to cheating detection (not using available statistical techniques, only acting on complaints for erasure analysis)
Throughout both administrations	Calls for genuine test integrity (e.g., current norms, data transparency, ethical practices)	Testing Companies & Complicit Academics	Continued use of outdated norms, refusal to provide data, legal threats against critics, development/promotion of "crib-sheet" test prep materials

Recommendations for Policy Changes to Foster True Testing Integrity, Drawing from Proposed Reforms

To move towards genuinely "honest testing," policymakers must implement comprehensive reforms that address both test design and administration. This includes adopting strict security policies, mandating the yearly rotation of equivalent test questions, requiring the use of independent external proctors, ensuring test booklets are sealed until administration, and shifting testing to the fall to reduce "teaching to the test" incentives.

Crucially, all students, including those in special education and bilingual programs, should be included in reported scores to provide an accurate picture of overall achievement. Technological methods, such as scanning for suspicious erasures and analyzing for cluster variance, should be routinely employed to detect irregularities.

Furthermore, accountability must extend to the testing industry itself: test publishers and scoring companies should be held legally liable for selling deceptive tests, using outdated norms, or willfully withholding evidence of testing irregularities. Finally, increased participation in and reliance on genuinely secure national assessments like NAEP, which are designed with robust

security and external oversight, is essential for providing accurate, comparable data on student achievement.

The persistent nature of the "Lake Wobegon" effect, despite decades of exposure and various reform attempts, highlights that the problem is not merely technical but deeply socio-political and economic. True reform necessitates addressing the fundamental incentives and power structures that benefit from inflated scores, rather than solely focusing on technical fixes to test instruments. Until these underlying motivations for resistance are confronted and realigned towards genuine accountability, the pursuit of honest testing will likely remain an uphill battle.

Works cited

1. Clinton Proposes National Exam, Again - Fairtest, <https://fairtest.org/article/clinton-proposes-national-exam-again/> 2. House OKs Compromise on Clinton National-Exam Plan - Los, <https://www.latimes.com/archives/la-xpm-1997-nov-08-mn-51547-story.html> 3. George W. Bush: Domestic Affairs | Miller Center, <https://millercenter.org/president/gwbush/domestic-affairs> 4. No Child Left Behind Act - Wikipedia, https://en.wikipedia.org/wiki/No_Child_Left_Behind_Act 5. Press Conference With President George W. Bush And Education Secretary Rod Paige To Introduce The President's Education Program, <https://georgewbush-whitehouse.archives.gov/news/releases/2001/01/20010123-2.html> 6. Does Texas' special education policy go back to George W. Bush's ..., <https://www.texastribune.org/2018/02/01/federal-assessment-cap-texas/> 7. A Tale of Two Approaches—The AFT, the NEA, and NCLB, <https://edpolicyinca.org/publications/tale-two-approaches-aft-nea-and-nclb> 8. NO CHILD LEFT BEHIND - American Federation of Teachers, <https://www.aft.org/resolution/no-child-left-behind>