

**The Accuracy and Effectiveness of  
Adequate Yearly Progress,  
NCLB's School Evaluation System**

by

**William J. Mathis  
Rutland Northeast Supervisory Union  
University of Vermont**

Education Policy Research Unit (EPRU)  
College of Education  
Division of Educational Leadership and Policy Studies  
Box 872411  
Arizona State University  
Tempe, AZ 85287-2411

**September 2006**

**EPRU | EDUCATION POLICY RESEARCH UNIT**

EPSL-0609-212-EPRU  
<http://edpolicylab.org>

This research was made possible by a grant from the Great Lakes Center for Education Research and Practice.

# **The Accuracy and Effectiveness of Adequate Yearly Progress, NCLB's School Evaluation System**

William J. Mathis

Rutland Northeast Supervisory Union

University of Vermont

## **Executive Summary**

Adequate Yearly Progress (AYP) is the key element of the accountability system mandated by the federal No Child Left Behind Act (NCLB). This report reveals that AYP in its 2006 form as the prime indicator of academic achievement is not supported by reliable evidence. Expecting all children to reach mastery level on their state's standardized tests by 2014, the fundamental requirement of AYP, is unrealistic. The growth model and other improvement proposals now on the table do not have sufficient power to resolve the underlying problems of the system. In addition, the program, whether conceived as implementation costs or remedial costs, is significantly underfunded in a way that will disproportionately penalize schools attended by the neediest children. Further, the curriculum is being narrowed to focus on tested areas at the cost of other vital educational purposes.

### **It is therefore recommended that:**

- AYP sanctions be suspended until the premises underlying them can be either confirmed or refuted by solid, scientific research and unintended, negative consequences can be avoided.

# **The Accuracy and Effectiveness of Adequate Yearly Progress, NCLB's School Evaluation System**

William J. Mathis

Rutland Northeast Supervisory Union

University of Vermont

## **Background**

Adequate Yearly Progress (AYP) is the linchpin of the federal No Child Left Behind Act (NCLB), passed in 2001 but only beginning to reach uniform full implementation in 2006. AYP uses year-to-year gains in standardized test scores as the primary measure of school quality and progress. Schools not making large enough gains—that is, not making Adequate Yearly Progress—are often labeled “failing”<sup>1</sup> in media reports. If a school does not meet AYP targets for several years, it faces increasing levels of sanctions, ultimately leading to restructuring, which may include a state takeover, conversion to a charter school, or hiring a private firm to manage the school. With such fundamental effects on schools and the future of public education, it is essential to understand whether the premises underlying this mechanism are valid, and whether evidence exists to support the claim that AYP can spark the reforms its supporters promise.

## ***What is Adequate Yearly Progress?***

The fundamental requirement of AYP is that all children meet mastery levels on their state's standardized tests by 2014. The tests must include annual reading and mathematics examinations in grades 3 through 8 and another state-selected basic skills examination in the high school grades, the latter frequently given in the 10th grade. By the 2007-2008 school year, each state must administer a science test in an elementary, middle, and high school grade. Requirements also include graduation rate as an "academic indicator" for high schools. For elementary schools, the most commonly used academic indicator is attendance rate.

Progress is measured by comparing the percentage of students who attain target scores in a given grade with the percent of students who attained them in the same grade the previous year. If the school meets its performance goals for the year (its annual measurable objective, or AMO), it is said to have made Adequate Yearly Progress. It is important to understand that this procedure compares two entirely different groups of students rather than measuring the progress (or lack thereof) of a particular cohort of students. Additionally, it does not reflect progress made by students who, despite having made strong gains, have fallen short of target scores.

The law requires states to have a uniform accountability system, applicable to all public schools (but not private schools), whether the school receives any federal money or not. The states also define goals for intermediate success (or AMOs) so that all students meet standards by 2014. For schools that are making significant progress but still falling short of targets, the NCLB accountability system includes a safety net called "safe harbor." If the school does not make AYP but does reduce the percentage of

children scoring below the mastery level by 10 percent per year, the school meets AYP through safe harbor. Since safe harbor often requires larger test score gains than AMOs, however, it is not applicable in many cases. For example, a school with 60 percent mastery in 2004 would be required to average only a 4 percent gain per year by 2014. For this same school, safe harbor's requirement of a 10 percent gain would not provide an effective alternate route.

AYP further requires that each subpopulation (such as minority groups, poor children, and special education students) also attain mastery by 2014. These groups generally start from a lower score level and, consequently, must make bigger improvements each year. Scores are "disaggregated" for each subgroup; that is, the score for each subgroup is considered separately. If a school fails to make the target growth for the entire group *or* for any one of the potential disaggregated groups, it is judged not to have attained AYP. For example, in Figure 1, this "non-performing school" did not reach proficiency in reading for students with disabilities (the red x). Although it passed all the other 36 criteria, this one deficiency means the school as a whole failed to make AYP.<sup>2</sup>

Figure 1

# THE NCSL TASK FORCE ON NO CHILD LEFT BEHIND

## Portrait of a Non-Performing School?

	Reading		Math		Other Indicator
	Participation	Proficiency	Participation	Proficiency	
All Students	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Economically Disadvantaged	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	Other indicator for secondary schools is the graduation rate. For elementary and middle schools, it is typically the attendance rate.  Additional indicator applies only to the school-wide population.
Asian/Pacific Islander	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	
Black	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	
Hispanic	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	
Native American	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	
White	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	
Students with Disabilities (SWD)	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	
Limited English Proficient (LEP) Students	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	

Source: National Conference of State Legislatures, used with permission.

## *AYP Sanctions*

NCLB requires states and districts to offer schools technical assistance in achieving AYP and allocates funds for doing so. The law requires that an improvement plan be developed with the assistance of teachers, principals, parents, higher education, and other parties the state agency determines appropriate.<sup>3</sup> Nonetheless, AYP is essentially a sanction-based system that increases the penalties on schools that fail to meet targets. After failing to make AYP for two years, a school is “identified” as “in need of improvement,” and it must set aside 20 percent of its Title I money to fund a school-choice option; if it fails for a third year, it must also fund mandated third-party supplemental education services (generally private tutoring). Failure to make AYP by a

school or district for two consecutive years triggers improvement-plan requirements and progressively increasing sanctions (see Table 1).<sup>4</sup> On May 17, 2006, U.S. Department of Education Secretary Spellings announced that states could apply to reverse the order of sanctions by allowing supplemental education services to be provided before school vouchers were required.<sup>5</sup> After the third year of failure to meet AYP goals, schools must formulate “school improvement” plans, which are put into effect at the end of that year if AYP goals are still not met. For each additional year that scores fall short of AYP targets, increasing sanctions are applied.



**Table 1: Comparison of NCLB Requirements for Identified Schools and Districts**

School Year After Being Identified (after not making AYP for two consecutive years)	Schools	Districts
<b>First Year</b>	<ul style="list-style-type: none"> <li>• Identified as in need of improvement</li> <li>• Develop improvement plan that addresses reasons for not making AYP</li> <li>• Offer students public school choice until school exits improvement</li> </ul>	<ul style="list-style-type: none"> <li>• Identified as in need of improvement</li> <li>• Develop improvement plan that addresses reasons for not making AYP</li> <li>• Can no longer be a direct provider of supplemental education services (tutoring) to its students</li> </ul>
<b>Second Year</b>	<ul style="list-style-type: none"> <li>• Implement school improvement plan</li> <li>• Continue to offer public school choice</li> <li>• Offer students supplemental education services until school exits improvement</li> <li>• By end of school year, district must implement corrective action, which may include replacing school staff, instituting new curriculum, decreasing management authority at school level, extending the school year or day, bringing in outside experts</li> </ul>	<ul style="list-style-type: none"> <li>• Implement district improvement plan</li> <li>• By end of school year, state must implement corrective action, which may include deferring program funds, instituting new curriculum, replacing district personnel, allowing students to attend school in another district, appointing new administrators, abolishing or restructuring the district</li> </ul>
<b>Third Year</b>	<ul style="list-style-type: none"> <li>• Continue to offer choice and supplemental education services</li> <li>• Implement corrective action</li> </ul>	<ul style="list-style-type: none"> <li>• Implement corrective action</li> </ul>
<b>Fourth Year</b>	<ul style="list-style-type: none"> <li>• Enter restructuring</li> <li>• Continue to offer choice and supplemental education services</li> <li>• District must develop and implement a 2-year plan which can include reopening the school as a charter school, making significant staff changes, turning school over to state education agency or private firm</li> </ul>	<ul style="list-style-type: none"> <li>• Implement corrective action</li> </ul>
<b>Fifth Year</b>	<ul style="list-style-type: none"> <li>• Implement school restructuring</li> <li>• Public school choice and supplemental education services must continue to be provided</li> </ul>	<ul style="list-style-type: none"> <li>• Implement corrective action</li> </ul>

Source: Center on Education Policy. Used with permission

## Recent Developments

The No Child Left Behind Act (NCLB) and the Adequate Yearly Progress (AYP) provisions have been the subject of considerable political controversy, the substance of which will be taken up in subsequent sections of this report. Public opinion on the act has been mixed. In the years since it was enacted, however, NCLB has met with official opposition from legislatures, governors, and/or state education officials in at least 31 of the 50 states. Opposition was reflected in laws, resolutions, or other official public calls for changes in the law.<sup>6</sup>

This official opposition is grounded in a variety of claims. One is that NCLB is unconstitutional and exceeds the authority of the federal government. In a 2005 report, the National Conference of State Legislatures (NCSL) cited the Tenth Amendment to the Constitution, which states that all powers not specifically delegated to the federal government are the domain of the states. Thus, they argued, education is a state responsibility, the funding provided by the federal government is disproportionate to the mandates, and the lack of money violates the express requirement of the law.<sup>7</sup>

This latter theme, the claim of a lack of adequate funding, has been a commonly expressed concern. Some groups that otherwise support the broad outlines and objectives of the law, such as the NAACP and the Education Trust, nonetheless have charged it lacks adequate funding.<sup>8</sup>

In early 2006, Secretary Spellings declared that the “rebellion” against NCLB had faded and that states were working toward compliance rather than mounting strenuous objections. At the same time, she announced a new “flexibility” toward implementing the law.<sup>9</sup> The most fundamental and visible proposed change is the use of the student

“growth model.” On May 17, 2006, Secretary Spellings also announced that North Carolina and Tennessee could use a growth model on an experimental basis for 2005-2006 with the possibility of six selected states being allowed to try their growth models for 2006-2007.<sup>10</sup> Budgetary support remains a major issue at the federal and state levels, with two major funding challenges winding through the courts. Finally, on the eve of the scheduled 2007 reauthorization of the law, a number of organizations have put forth their own recommendations for improving the AYP system.

### ***Public Support***

A variety of public polls by Phi Delta Kappa, *Education Week*, and the Public Education Network (PEN) indicate that the general public knows little about the law. As poll respondents are told more about its details, however, their negative response rate increases.

The annual Phi Delta Kappa/Gallup Poll on education released in fall 2005 found that only 40 percent of the public said they knew “a great deal” or “a fair amount” about NCLB.<sup>11</sup> On individual details that make up the act’s approach, majorities of the public appear to be opposed. For example, 68 percent of respondents said they disagreed with relying on a single test score to identify whether a school needed improvement. Fully 90 percent of respondents said it was “very important” to close the achievement gap between white and minority students, but a larger majority appeared skeptical about using the threat of transferring children to other schools as a means of achieving that end. Asked how they would respond if their own child was attending a school labeled “in need of improvement,” 79 percent said they would prefer to see “additional efforts made in the child’s present school” rather than acting to transfer the child to another school. In an

open-ended question format, the PDK poll asks respondents to name “the biggest problem” faced by public schools in their communities. “Lack of financial support” was selected as the “biggest problem” by the largest plurality of PDK poll respondents—20 percent, nearly twice the percentage who chose the second- and third-highest responses (“overcrowded schools,” 11%, and “lack of discipline,” 10%).<sup>12</sup> Other surveys have shown strong support for the law’s purposes but dissatisfaction with the notion of punishments and an overemphasis on testing.<sup>13</sup>

### ***Flexibility***

What federal officials describe as flexibility has taken several forms. These include allowing more testing exemptions for a small portion of special education students under certain circumstances; relaxing testing rules for English Language Learners; changing minimum group sizes and confidence intervals; providing relief for Hurricane Katrina affected areas; allowing a pilot project to change the order of sanctions; and using a system based on how much the same students learn over time (a “growth model”).<sup>14</sup> Critics of the law have not been enthusiastic about such flexibility. Some have argued that the proposals are inadequate to resolve the problems.<sup>15</sup> A Harvard Civil Rights Project report, meanwhile, argues that the flexibility allows school districts to distort or evade the purpose of the law without addressing the perceived shortcomings of the AYP system.<sup>16</sup>

### ***The Growth Model***

The most prominent proposed change in AYP sought by many states and educational organizations is implementation of a growth, or value-added, model. The

current AYP system compares this year's students with last year's students in the same grade, a procedure known as the "status model." The Department of Education in early 2006 released guidelines to allow a small number of states to try this approach provided they met seven core principles, including having implemented testing in all required grades and subject areas by 2004-2005, and having a student data tracking system in place. The trial states will still be expected to have universal proficiency by 2014. Twenty states submitted proposals to the Department of Education in February 2006 even though many did not meet the eligibility criteria. Recommendations for approval went to the Secretary in May 2006.<sup>17</sup> As earlier noted, Tennessee and North Carolina have been approved with a new review scheduled for six states in fall 2006.

### ***The Budget Outlook***

While Connecticut, local districts and the National Education Association (NEA) were bringing two separate lawsuits for inadequate funding, the appropriations for fiscal year 2006 were cut 1 percent across the board. For Fiscal Year 2007, the President has proposed cutting the education budget by another \$2.1 billion or nearly four percent.<sup>18</sup> In the Fiscal Year 2007 proposal, Title I, the primary federal source of remedial funds, is level funded, and special education is recommended for only nominal increases. Numerous other related titles and programs (such as teacher professional development and technology) also are cut or eliminated in the President's plan. It must be noted that Congress typically increases these numbers during the budget process.

The President's announcement came at the same time that the generally pro-NCLB Education Trust reported that low-income and minority students are short-changed by and within the states. In 2003 dollars, the Trust reports that poor children receive an

average of \$1,436 less each year than children in wealthy districts, a gap that has increased since 1997. This is of particular concern, given that the schools attended by poor children historically have lower standardized test scores and consequently are those least likely to achieve AYP.<sup>19</sup>

Inadequate funding has arguably been the single largest objection to NCLB and AYP. With new administrative costs adding 2 percent to school budgets and remedial costs adding an additional 27.5 percent, the President's proposed cuts run counter to what critics assert is a fundamental shortcoming of NCLB.<sup>20</sup>

### ***Calls for the Reform of AYP***

A number of organizations have called for changes including the National School Board Association (NSBA), the NEA, a group of Chief State School Officers, the NCSL, and the Center on Education Policy (CEP). Proposed improvements from these groups range from mild changes to the *de facto* elimination of the AYP system. In October 2004, the NSBA recommended that the AYP system be determined by individual states, that a continuous scale of progress be used, that no one group's scores cause the entire school to fail, and that growth expectations be flexible, depending on where a school started and what obstacles it faces.<sup>21</sup> The NEA put forth statistical projections in July 2004 from a number of states and concluded that the AYP system is so fundamentally flawed that it cannot work.<sup>22</sup> Recommendations include using additional indicators to supplement test results, increasing flexibility, and changing sanctions.

A group of sixteen state education superintendents and commissioners (Chief State School Officers) said in March 2004 that the AYP requirements could not reasonably be met under current interpretations and recommended that states be given the

flexibility to determine and use their own definitions and programs for determining adequate progress.<sup>23</sup> Similarly, the NCSL noted in February 2005 that AYP is fundamentally flawed and recommended that states be allowed to develop their own accountability systems. It also has endorsed student growth models and, like the NSBA, recommended eliminating the provision for labeling an entire school “failing” because only one of its 37 disaggregated groups failed. NCSL has pointed out that the law is underfunded, that federal influence is disproportionate to the small federal investment, and that the constitutionality of NCLB is questionable.<sup>24</sup>

CEP, an independent think tank that has closely studied the implementation of NCLB, held a conference in 2004 with the purpose of defining improvements to the law. The attendees generally subscribed to the view that it was impossible to meet AYP requirements and offered several recommendations: setting attainable performance goals, giving credit for growth in scores, allowing the states flexibility to design their own accountability systems, and using measures beyond test scores in evaluating schools.<sup>25</sup>

Common threads among these diverse panels are reflected in the joint statement of 30 educational organizations.<sup>26</sup> They concurred that AYP was unrealistic and unworkable, and that NCLB was dramatically underfunded. They advocated for more inclusive measures of school performance, the use of growth scores, and allowing states to design their own accountability systems and consequences.

## **Available Data**

While the No Child Left Behind Act (NCLB) requires scientifically-based evidence to support proposed education reforms, there is as yet little scientific data on the efficacy of the Adequate Yearly Progress (AYP) mandate itself.<sup>27</sup> Evidence that bears on

the subject can be sorted into one of a number of questions, some of which address AYP specifically and others of which are more broadly focused. Those questions include:

1. How effective are high-stakes standardized tests in assessing student achievement and school quality?
2. Have the AYP requirements prompted increases in test scores in public schools?
3. Are schools with poor and diverse populations identified more frequently as not making AYP?
4. Is curriculum being narrowed to the tested subjects, to the detriment of the larger educational mission?
5. Is AYP adequately funded?
6. What do existing growth rates or other data suggest about whether AYP requirements are realistic and achievable?



## *How Effective are High-Stakes Standardized Tests in Assessing Student Achievement and School Quality?*

Certainly many volumes have been written on this question alone. Nonetheless, it merits a review in light of the critical role high-stakes standardized testing plays in measuring AYP. Multiple controversies abound in this hotly debated area. In this section, three issues are discussed: 1) test validity, 2) the relative merits of the “status model” and the “growth model,” and 3) the record of high-stakes testing as a means of closing the achievement gap.

### **Test Validity: “Alignment”**

Validity is whether the test measures what it claims to measure. In the AYP regime, validity has been defined essentially as content validity. That is, whether the tests are aligned with state standards in proportionate measure and difficulty. The precision of that alignment, however, is the subject of much dispute.<sup>28</sup> Tests are most commonly designed by groups of subject matter specialists who compare the test content objectives with the state standards. However, the match between test content and learning objectives has shown considerable variability based on the set of raters making the judgments.<sup>29</sup> Additionally, statewide tests generally address lower-order content rather than the more abstract and higher-level requirements of that state’s own standards.<sup>30</sup>

Further, as illustrated by extensive controversies over the best way to teach reading, experts disagree among themselves about the best way to organize and measure subject matter content. Thus, test performance may hinge directly on a particular set of

instructional or philosophical assumptions which may not be explicit nor reflect the practices or consensus of the field.<sup>31</sup>

## **The Status Model vs. The Growth Model**

As has been noted, the current AYP system uses the “status model,” which compares different cohorts of students. That is, this year’s fifth graders’ test scores are compared with those of last year’s fifth graders, and so on, to measure the school’s progress. Research suggests that such comparisons, however, introduce significant error and fail to properly identify school effects on learning. Kane and Staiger,<sup>32</sup> for example, found that comparisons between different cohorts could result in 70 percent error.

The desired alternative of many is to compare the same students with themselves over time, indicating their growth (hence the term “growth model”). For example, students’ fifth grade scores would be compared with their scores as sixth graders. This approach would allow consideration of growth even if all the growth was still below the mastery level. Advocates of growth models believe this technique would also better measure the effects of schools (the “value-added” to a child’s performance), particularly in places where achievement scores are highly influenced by out-of-school factors such as poverty.<sup>33</sup>

A body of evidence, however, suggests limitations in the growth model as well, specifically regarding the ability to compare different tests across different grade levels.<sup>34</sup> Commonly used test construction models assume that knowledge progresses steadily up through the grade levels. Test items are arranged on a one-dimensional scale based on the assumption that students acquire knowledge in a linear (or latent trait) fashion. Once new content is taught, a previously difficult item becomes an easy item and the item’s

weight changes. The result, however, is that test scores can be thrown out of alignment from year to year. In New York's case, this phenomenon caused the passing score to be established inappropriately high; consequently, the number of correct responses necessary to earn a passing score was cut to 36 items from the original 51 items.<sup>35</sup>

The primary alternative approach is to define mastery within one grade level and then compare it with the next grade's mastery expectation. Scores from one grade to the next are compared to see if growth has been made. The two tests, however, now measure different things; fourth grade math test scores that measure division may be subtracted from fifth grade scores that measure geometry. The result led Tucker, examining the New York troubles, to conclude: "The reality is that passing scores on standards-based tests over time are very unlikely to be comparable at the level of precision that justifies high-stakes consequences . . . ."<sup>36</sup>

## **High Stakes Tests and the Achievement Gap**

A body of research that goes beyond the testing regime implemented under NCLB raises questions about the ability of such high-stakes assessment to raise achievement among lower-achieving students. Harris and Herrington examined the last half-century of accountability-based reforms and found these approaches have only a moderate influence on achievement. They conclude that high-stakes assessment does not hold sufficient promise for closing the achievement gap.<sup>37</sup> Similarly, in comparing 25 states with high- and low-pressure systems of state accountability, Nichols, *et al.* did not find evidence that high-stakes systems such as AYP improved state National Assessment of Educational Progress (NAEP) scores. They did find that high-pressure systems resulted in more dropouts and lower retention rates.<sup>38</sup> In looking at the various studies of the

effects of high-stakes accountability on test scores, Ritter found that the studies were split, with just over half showing positive results and the others showing mixed or negative results. He concludes that the evaluation data on these efforts are “woefully incomplete.”<sup>39</sup> Further confounding matters, Starr has speculated that lower NAEP scores may reflect greater disinterest in that particular exam on the part of students because of the NAEP’s low stakes.<sup>40</sup>

### ***Achievement Test Scores: Have AYP Requirements Prompted Increases in Test Scores in Public Schools?***

Various studies have reported recent gains in achievement test results, particularly in the lower grades and in math. Despite federal claims of NCLB success, however, NAEP gains cannot be ascribed to AYP, to NCLB, or to any other particular cause with any degree of assurance.<sup>41</sup> The primary reason is that the reported score increases are correlational in nature and the growth claim dates do not match the NCLB implementation dates.

For the four reported NAEP ethnic groups (African-American, Hispanics, Whites and Asians), NAEP scores in reading and math have improved from 1984-2005 (NCLB was signed into law in 2002).<sup>42</sup> The greatest gains have been in mathematics. Overall, however, test score improvements have been modest. This seemingly inconsistent result reflects the increase, over the same time period, in the proportion of African American and Hispanic students in the tested population. Since the test scores of these groups have lower starting points, their increasing numbers brought the overall average down, even though all groups improved significantly.<sup>43</sup>

The Education Trust, which supports NCLB and the AYP requirements, found that math proficiency scores improved in 23 of 24 states. Reading scores improved but not by as large a margin as math scores. In a later study, the Trust did not find strong middle and high school gains from 2003 to 2005.<sup>44</sup>

In late 2005, Secretary Spellings asserted that recent NAEP test score improvements were “proof that No Child Left Behind is working.”<sup>45</sup> That conclusion is not supported by the evidence. The claim of increased test scores as a result of NCLB is associational. Causality cannot be assumed from this relationship. Without more rigorous experimental design, the reason or reasons for improvements remain unexplained. As Koretz has noted, any gains may be the result of pre-existing state programs or some other set of circumstances.<sup>46</sup>

A further reason for skepticism is that the law’s implementation, beginning in 2002, was slow and turbulent. NCLB-driven changes impacting classroom practices would not likely move into classrooms until 2003 or 2004 which would limit the potential causal effects of the law. It seems unlikely, therefore, that NAEP gains shown by 2005 would be attributable to NCLB and AYP. The Northwest Evaluation Association, using its own individual student data base of 320,000 children, reported test score increases from 2003 to 2005. A close look at the data, however, shows that the pre-NCLB test score increases (2000 to 2003) were greater than those from the post-NCLB 2003 to 2005 years.<sup>47</sup> Without more detailed research, therefore, one might as plausibly conclude that NCLB actually put a damper on increases. Of course, such a conclusion is also correlational and not supported by research, either.

Linn has pointed out another conflicting factor: High stakes assessment efforts typically show immediate score declines with the introduction of a new test, then scores characteristically show gains in early years before leveling-off.<sup>48</sup> Thus, it is too soon to know if the improvements are a spike or a trend, regardless of the cause. In any interpretation, the recent noted gains in test scores would not be sufficient to reach the mandated outcomes by 2013-2014.

### ***Trends in Identified Schools: Are Poor and Diverse Schools Identified More Frequently?***

Poverty has been shown to account for almost half the variation in test scores among schools.<sup>49</sup> Poor schools with lower test scores are identified as not making AYP more frequently than affluent schools.<sup>50</sup> As students attending poorer schools generally start with much lower scores than their more affluent counterparts, they must make much greater AYP advances over the same amount of time. Since a school must make AYP for every sub-population, schools with more ethnic groups have more opportunities to fail than homogenous schools.<sup>51</sup> The result appears likely to lead to disproportionate identification of racially diverse schools as not making AYP.<sup>52</sup> Proponents of the system do not deny this likely outcome, but argue that the only way to help ethnically diverse schools is to continue to shine a bright and clear light on those that fail to meet test-based standards for all groups.<sup>53</sup>

The Center on Education Policy found that identified schools have become more concentrated in large urban school districts, and in schools that are ethnically diverse.<sup>54</sup>

***Is curriculum being narrowed to the tested subjects, to the detriment of the larger educational mission?***

There is considerable evidence that high-stakes tests generally lead to a narrowing of the curriculum and push instruction toward lower-order cognitive skills.<sup>55</sup> Using Texas as an example, Darling-Hammond demonstrated that a rise in basic skill scores on the lower-order state tests was not accompanied by a rise in more comprehensive NAEP test scores.<sup>56</sup>

The Center for Education Policy's (CEP) Year 3 national sample found 27 percent of schools said they reduced their social studies instruction and 22 percent their science instruction due to NCLB curricular demands.<sup>57</sup> By the Year 4 CEP report, the proportions had increased—71 percent of districts reported curriculum narrowing to the tested subjects.<sup>58</sup> The Rockefeller Center at Dartmouth College reported that 70 percent of New Hampshire and 83 percent of Vermont Superintendents said that tested subject matters are receiving more attention at the expense of social studies, arts, and science.<sup>59</sup> There is evidence as well that curriculum narrowing falls disproportionately along socio-economic lines. Hargreaves, in his studies of accountability in Ontario and New York, states that students in more affluent and high-scoring schools continue to receive a diverse and rich program, while urban and minority school students there are subjected to a micro-managed regimen of drill and practice in basic skills areas.<sup>60</sup>

Other observers have challenged these contentions, however. In 2005, Kati Haycock, director of the Education Trust, a school-reform advocacy group that supports NCLB, testified before Congress that high-performing schools did not narrow the curriculum: "I have never come across a high-performing school that was inordinately

focused on ‘drill and kill’ or test-prep strategies. High-poverty schools where learners are excelling tend to be the most dynamic, creative, engaging learning environments.”<sup>61</sup>

### ***Is AYP Funded Adequately to Achieve Its Purposes?***

Perhaps the most repeated concern about NCLB is the lack of adequate funding. State governments as well as critics and advocates of the law have claimed NCLB is dramatically underfunded. The bipartisan National Conference of State Legislatures found the law underfunded both for carrying-out new administrative burdens as well as inadequately funded for providing necessary remediation programs.<sup>62</sup> The independent Center on Education Policy reported two-thirds of state agencies said they were underfunded in meeting administrative burdens and 80 percent of districts reported similar underfunding.<sup>63</sup> The pro-NCLB Education Trust finds that low-income and minority children are short-changed and that unequal resources means unequal opportunities. The Trust, however, sees these inequalities as a state problem rather than a federal problem.<sup>64</sup>

The Bush Administration and its allies, defending the current funding as adequate, cited the following as evidence:

- Former Secretary of Education Rod Paige asserted—contrary to claimed full-funding commitments to legislators—that there was never any such commitment for full funding.<sup>65</sup>
- Former Secretary Paige also asserted that the budget had increased by 51 percent from 2001 to 2003.<sup>66</sup>



- Rep. John Boehner (R-Ohio) in 2004 contended NCLB and AYP were adequately funded because appropriations were not fully spent each year—that is, there was “money left over.”<sup>67</sup>

These defenses of current funding levels, however, do not address the underlying criticisms. For example, the size of the growth in funding does not address the actual cost of administering new programs. The argument that money is “left over,” meanwhile, ignores federal policies that encouraged states and districts to “carry over” money across several years. Additionally, the unused residual has been calculated at 0.03 percent – that is, 3/100<sup>ths</sup> of one percent – of total education spending.<sup>68</sup>

### **Alternative Measures of Funding Adequacy**

Under other approaches to analyzing funding levels, there is evidence that the law is not adequately funded:

**Statutory Funding Level:** The law says that for each student counted as socially or economically deprived an additional 40 percent of the average statewide per pupil expenditure (within caps for very high and very low spending states) will be appropriated. Based on this definition of funding, NCLB is funded at only 41 percent of the figure required by law. This discrepancy has not been discussed extensively in funding debates.<sup>69</sup>

In 2005 and 2006, the cost debate parameters shifted toward the actual hard costs of implementing NCLB systems and the costs of remediation or the cost of bringing all children up to standards.

**New Administrative “Hard” Costs:** “Hard costs” are defined as the costs to build, implement and score new tests, develop AYP data systems, and support technical assistance capabilities. A group of new studies, primarily under the aegis of the Council of Chief State School Officers, argue that new hard costs increase total education spending about 2 to 2-1/2 percent per student. If the total increase in new federal NCLB funds were applied to the hard costs, they would cover only about one-third to one-half of these costs.<sup>70</sup> New appropriations for Title I, the primary AYP affected program, have increased from \$8 billion to \$12 billion, or less than one percent of total education spending since 2001. Federal funds were cut in 2006, however, and President Bush has requested further cuts in 2007.<sup>71</sup>

Claiming NCLB is adequately funded, Peyser and Costrell contend that states should have already built these systems and so the costs should not be ascribed to NCLB or AYP. In their view, compensatory and special education money can be “reassigned” to AYP. Therefore, the NCLB system is adequately funded.<sup>72</sup>

**Costs of Remediation:** A wide variety of factors that are beyond the control of schools, including but not limited to poor nutrition, poor health, less stable households, and lack of parental support for education, all contribute to the higher number of learning difficulties for children living in impacted poverty. Researchers have found that the costs of providing necessary services for children to overcome learning deficits that are a consequence of living in poverty are approximately twice<sup>73</sup> the current cost of education. This author’s review of almost 50 state-level studies of what would be required to provide an adequate education (that is, all children meeting standards) conducted by a wide variety of independent researchers and sponsored by a similarly wide range of

groups, found that the median increase in expenditures was 27.5 percent, or \$137.8 billion, in new remedial costs alone.<sup>74</sup>

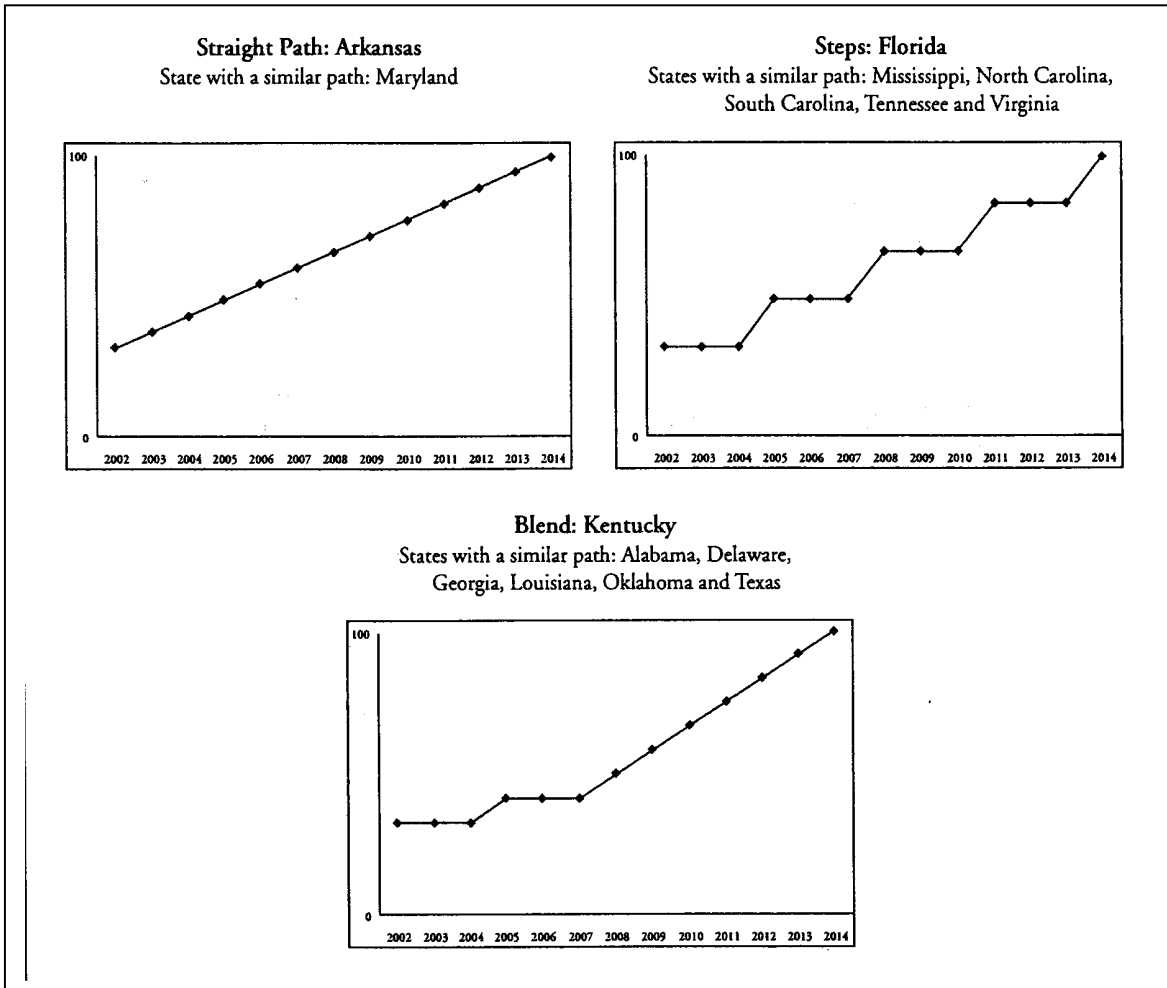
These assertions of under funding are countered by claims that the way in which the cost estimates are produced are alchemy,<sup>75</sup> that schools have enough money but they are using it inefficiently,<sup>76</sup> and that the costs for educating poor children may be higher but that is not a federal responsibility.<sup>77</sup>

***What do existing growth rates or other data suggest about whether AYP requirements are realistic and achievable?***

State attempts to measure schools' AYP generally employ one of three basic roadmaps, although there is variation in detail (see Figure 2). Linear, or Straight Line, assumes regular, incremental growth of the same amount each year. Performance targets are set for the group as a whole and for each subgroup. The Stair-Step model, used most frequently by states, assumes that test score growth will occur in incremental steps (see Florida graph in Figure 2). The assumption is that improved instruction will improve results, but it takes a period of years before the gains appear in test-scores; then, a period of new program improvements is introduced, gestates, becomes institutionalized, and results in a new leap forward. Then, the cycle is repeated. In a typical case, this model requires increases every three years. Finally, the Blended, Balloon, or Accelerated Curve model assumes that several years of teacher preparation, program phase-ins, and implementation are needed before improvement is evident. Thus, small gains are expected in the early years and the rate of progress ascends rapidly upward as the new improvements take hold, producing total proficiency by the 2014 deadline.

Figure 2

States Chart Different Paths to the *No Child Left Behind* Goal of 100 Percent Student Proficiency by 2014



Source: Prepared by Southern Regional Educational Board, April 2005, used by permission

Wiley, *et. al.* modeled AYP results in six states in May 2005 and found that virtually all schools will be identified on or before 2014.<sup>78</sup> Regardless of the AMO model used, the percentage of students having to demonstrate mastery increases continuously to the required 100 percent level in 2014 (with some small exceptions).

Policy analysts, such as Lee's analyses in Maine and Kentucky, have done similar projections in many states.<sup>79</sup>

In the early years, the increase in identified schools has not happened as rapidly as predicted. The percentage of identified schools has remained steady at about 13 percent of schools for the two years of available data.<sup>80</sup> States have received federal approval to change test confidence intervals, performance indices, minimum subgroup sizes, and retesting procedures. They also have implemented new testing programs and changed their mastery cut-off scores. The sum effect of all these factors has been a leveling-off of the percentage of identified schools to date.<sup>81</sup> As these changes work through the system, however, it seems likely that the number of identified schools will resume rising as the AMO expectations increase at a faster rate than student achievement scores.

The Associated Press did their own analyses of "uncounted" students which they released in April 2006. They estimated 1.9 million children's scores were not being reported (even though these scores were reported in the school totals) due to the minimum cell size restrictions adopted by states and approved by the federal government. The objective of minimum subgroup sizes in reporting disaggregated scores is to assure statistical reliability of the results. This media analysis has prompted Secretary Spellings to say she will close the "loophole" and Congressional hearings will be held. If, as a result, minimum cell sizes are reduced, the effect will be to identify more schools as not making AYP and at a faster rate.<sup>82</sup>

Although not a scientific, peer-reviewed study, an Associated Press report in May 2006 offered indications that the anticipated increase in identified schools may already be happening. The article said that 1,750 schools – 3 percent of the nation's 53,000 schools

receiving federal funds – had been ordered into “restructuring.” Michael Petrilli of the Thomas B. Fordham Foundation, a supporter of the law’s sanctions, predicted that “It’s just a matter of time before we see upwards of 10,000 schools in restructuring”—more than 1 in 5.<sup>83</sup>

Supporters of NCLB’s sanctions regime offer evidence that the AYP goals are possible in the form of reports of “excellent” schools (as defined by test score gains) in deprived neighborhoods that have seemingly beaten the odds.<sup>84</sup> Skeptics counter with evidence indicating that many such so-called “excellent” schools are statistical outliers that do not maintain such progress over years.<sup>85</sup> Harris in 2006 examined data underlying two reports that purported to describe successful, high-poverty, high-minority schools. He found that “the number of high-flying schools is significantly smaller” than either report indicated. Only about 1.1 percent of the schools’ claimed to be high fliers could maintain altitude.<sup>86</sup>

## **Discussion and Analysis of Available Data**

Taken together, the evidence that is available on AYP suggests the following conclusions:

- There is near universal consensus that test score gains that may be attributable to the AYP process are modest and insufficient to achieve the goal of 100 percent proficiency by 2014.
- There are insufficient funds to effectively implement the program. Calls for adequate funding of NCLB come from the public, legislators of both

parties, and advocacy groups ranging from the Education Trust to the National Education Association.

- Curriculum is being narrowed to focus on tested areas at the cost of other vital educational purposes.

Many groups have weighed in on suggested improvements to the AYP system. The changes proposed to date appear to offer little, if any, solutions for the more critical underlying problems.<sup>87</sup> For example, there are strong endorsements of growth models and other “flexibility” measures. Since all students must still meet the same standards by 2014 under a growth model and no appreciable increase in resources is on the horizon, the underlying problems are not resolved. If the minimum size for disaggregated groups is reduced, the sum effect would be to make an already unworkable system less workable.

The word “flexibility” has been a vague codeword for reforming NCLB. For Secretary Spellings, this means a trial of growth models, but for the group of chiefs petitioning the secretary, flexibility means replacing the national AYP system with state accountability systems. Both groups use the same word but the meaning is significantly different.

Since 2001, the debate on the costs of NCLB has shifted toward (1) examining the hard costs of implementing the program such as the cost of assessment and accountability systems, and (2) the costs of providing the extensive social and educational programs necessary for children living in impacted poverty to succeed in school. The hard cost studies indicate that new federal appropriations do not cover the implementation costs. The median of almost fifty state level remediation cost studies show an additional increase of 27.5 percent is needed on top of new administrative costs. The discrepancy

between the funding provided and estimates of what may be necessary to achieve learning goals, particularly in high-poverty schools, may suggest an opening for both new legislative action as well as litigation at the state and federal levels. Assuming new litigation occurs (and is successful), the time consumed in such litigation and resulting legislative battles would not result in resources reaching children in a fashion that would allow the 2014 deadline to be reached.

While Hanushek and others have questioned the way in which remediation costs are calculated,<sup>88</sup> they have yet to provide a superior costing mechanism to the adequacy studies or to the current system of political determination of education aid.

“Inefficiencies” as defined by Taylor and others rest on the assumption that spending above the average, regardless of differences in circumstances is inefficient and such monies could be recouped.<sup>89</sup> Given the politics and nature of spending variations, such an assumption is unlikely to be sustained or any funds recouped. Thus, the weight of the adequacy studies indicating significant new costs is not contradicted.

While the requirement to disaggregate scores by ethnic and socio-economic groups may cast a new spotlight on the national problem of the educational disadvantages of poor and minority children,<sup>90</sup> there is as yet, no evidence that, by itself, AYP demands will adequately provide schools serving children in poverty with the facilities, the learning resources, the qualified staff, or community support services. Schools in neighborhoods of impacted poverty will not reach AYP and sustain gains by intensely implementing a phonics program or testing students more frequently. Among many other needs, child care, early education, balanced nutrition, after-school opportunities, summer



programs, and social opportunities are essential for the massive reculturing fundamental to no child being left behind.<sup>91</sup> This is neither cost-free nor free of controversy.

The coupling of high-stakes consequences with test scores increasingly appears to narrow curriculum in the opinion of state level and local level educators. Such a shift in curriculum has implications for the decline of the traditional liberal arts curriculum as areas not so easily measured – such as art, physical education, and civic participation – become overshadowed by the high visibility and intensity of test-based sanctions. If education is redefined as clearing progressively higher hurdles on test scores, then the curriculum will naturally face incentives to conform to the content, pace, and timing of the testing cycle. Further, the schools with the lowest scores, which are most often found in poor areas and inner cities, will run afoul of AYP sanctions faster and, consequently, narrow the curriculum at a quicker rate for the neediest children.<sup>92</sup> For society, this brings into focus the philosophical question of whether schools are to define their aims by test scores of cognate or whether they also must teach the values, attitudes, and knowledge necessary in a democratic society.

## **Recommendations**

It is recommended that:

- AYP sanctions be suspended until the premises underlying them can be either confirmed or refuted by solid, scientific research and unintended, negative consequences can be avoided.

## Notes & References

<sup>1</sup> A sampling of “failing schools” media reports include:

Nevada governor wants \$100 million for failing schools. (2006, March 6). *Las Vegas Sun*. Retrieved March 13, 2006, from <http://www.lasvegassun.com/sunbin/stories/nevada/2006/mar/06/030610891.html>

A bold plan for failing schools. (2006, March 13). *The American Daily*, Phoenix. Retrieved March 13, 2006, from <http://americandaily.com/article/12100>

Jones, H.V. (2006, March 13). Some balance needed to remedy Bush’s imperial regime. *Fredericksburg (VA) Free Lance-Star*. Retrieved March 13, 2006, from [http://fredericksburg.com/News/FLS/2006/032006/03132006/printer\\_friendly](http://fredericksburg.com/News/FLS/2006/032006/03132006/printer_friendly)

<sup>2</sup> National Conference of State Legislatures (2005, February). *Task force on no child left behind: final report*. Denver CO: Author. p.18

<sup>3</sup> Public Law 107-110-Jan. 8, 2002 115 Stat. 1499.

<sup>4</sup> Center on Education Policy. (2005, March). *Identifying school districts for improvement and corrective action under the no child left behind act*. Washington, DC: CEP

<sup>5</sup> U. S. Department of Education. (2006, May 17). *Secretary Spellings approves Tennessee and North Carolina growth model pilots for 2005-2006* [Press release].

<sup>6</sup> Center on Education Policy. (2005, March). *From the capital to the classroom: Year 3 of the No Child Left Behind Act*. Washington, DC: Author. p. 58.

Mathis, W.J., Fleming, K., & Lecker, W. (2005, March). *The NCLB rebellion: Money, politics and student learning*. Paper presented at the American Education Finance Association’s annual meeting. Louisville, KY.

Young, S. (2005, May 31). *No Child Left Behind: 2005 legislative activity*. Denver, CO: National Conference of State Legislatures.

National Conference of State Legislatures. (2005, February). *Task force on no child left behind: Final report*. Denver: Author.

<sup>7</sup> Young, S. (2005, May 31). *No Child Left Behind: 2005 legislative activity*. Denver: National Conference of State Legislatures.

National Conference of State Legislatures. (2005, February). *Task force on No Child Left Behind: Final report*. Denver, CO: Author.

<sup>8</sup> See, for example:

“Education Department,” (undated). NAACP. Retrieved May 11, 2006, from [http://www.naACP.org/rat/education/education\\_index.html](http://www.naACP.org/rat/education/education_index.html)

See also

The Education Trust (Winter 2005). *The funding gap 2005. A special report of the Education Trust*. Washington, DC: Author.

Joint organizational statement on the No Child Left Behind (NCLB) Act. (2004, October 21). Retrieved May 11, 2006, from <http://www.schoolfunding.info/news/federal/10-28-04jointstatement.php3>

<sup>9</sup> Davis, Michelle R. (2006, January 18). Yearly progress. *Education Week*, 25(19), pp.30-32.

<sup>10</sup> U. S. Department of Education (2006, May 17). *Secretary Spellings approves Tennessee and North Carolina growth model pilots for 2005-2006* [Press release].

- <sup>11</sup> Rose, L. C. and Gallup, A. M. (2005, September). The 37<sup>th</sup> annual Phi Delta Kappa/Gallup poll of the public's attitudes toward the public schools. *Phi Delta Kappan*, 87(3), 41-58. Retrieved May 11, 2006, from <http://www.pdkintl.org/kappan/k0509pol.htm>
- <sup>12</sup> Rose, L. C. and Gallup, A. M. (2005, September). The 37<sup>th</sup> annual Phi Delta Kappa/Gallup poll of the public's attitudes toward the public schools. *Phi Delta Kappan*, 87(3), pp. 41-58. Retrieved May 11, 2006, from <http://www.pdkintl.org/kappan/k0509pol.htm>
- <sup>13</sup> National Education Association (2004, January 14). *Bipartisan education poll finds need for more flexibility and funding in NCLB* [Press release].
- Robelon, E. (2004, April 7). Opposition to school law growing, poll says. *Education Week*, 23(30), pp. 22, 27.
- Goodwin, B. (2003, July). Digging deeper: Where does the public stand on standards-based education? Issues Brief. Aurora, CO: Mid-Continent Research for Education and Learning.
- <sup>14</sup> Center on Education Policy (2005). *States test limits of federal AYP flexibility*. Retrieved on December 9, 2005, from <http://www.cep-dc.org/federalprograms/federal/AYPflexibility/stateaccountabilityPlanReport.doc>
- <sup>15</sup> Linn, R.L. (2004, July 28). *Rethinking the no child left behind accountability system*. Paper prepared for the forum on ideas to improve the NCLB accountability Provisions Washington: Center on Education Policy.
- Linn, R.L., Baker, E.L., & Betebenner, D.W. (2002, September). Accountability systems: Implications of requirements of the No Child Left Behind Act of 2001. *Educational Researcher*, 31(6), 3-16.
- <sup>16</sup> Sunderman, G.L. (2006, February). *The unraveling of No Child Left Behind: How negotiated changes transform the law*. Cambridge MA.: The Civil Rights Project, Harvard University.
- Olson, L. (2005, December 14). Room to maneuver. A progress report on the No Child Left Behind Act. *Education Week*, 25(15). A special pullout section.
- <sup>17</sup> U.S. Department of Education. (2006, January 25). Peer review guidance for the NCLB growth model pilot application. Washington DC: Author. Retrieved January 28, 2006, from <http://www.ed.gov/policy/elsec/guid/growthmodelguidance.pdf>
- <sup>18</sup> Alliance for Excellent Education. (2006, February 7). President Bush releases fiscal year 2007 budget. *Straight A's*, 6(3), pp 1-8.
- <sup>19</sup> The Education Trust (Winter 2005). *The funding gap 2005. A special report of the Education Trust*. Washington, DC: Author.
- <sup>20</sup> Palaich, R. (Personal email communication regarding the CCSSO NCLB administrative cost studies, February 28, 2006).
- Mathis, W. J. (2005). The cost of implementing the federal No Child Left Behind Act: Different assumptions, different answers. *Peabody Journal of Education*, 80(2), .91-120.
- <sup>21</sup> National School Boards Association. (2004, October). *Ensuring the goals of No Child Left Behind: Recommendations for legislative and regulatory change*. Alexandria, VA.: Author.
- <sup>22</sup> Packer, J. (2004, July 28). *No Child Left Behind (NCLB): Fundamental flaws need fixing*. Paper presented at the Center on Educational Policy's Forum on Ideas to Improve the Accountability Provisions Under the No Child Left Behind Act. Washington, DC: National Education Association.
- <sup>23</sup> Education Officials for NCLB Flexibility (2004, March 24). Letter to Secretary R. R. Paige (no address).
- <sup>24</sup> National Conference of State Legislatures. (2005, February). *Task force on No Child Left Behind: Final report*. Denver, CO: Author.

- <sup>25</sup> Center on Education Policy (2004, September). My school didn't make adequate yearly progress – so what does that mean? *Testtalk for Leaders*. Issue 3. Washington: Author.
- <sup>26</sup> Joint organizational statement on the No Child Left Behind (NCLB) Act. (2004, October 21). Thirty signatory groups including the National School Boards Association, the American Association of School Administrators and the National Education Association. Retrieved May 11, 2006, from <http://www.schoolfunding.info/news/federal/10-28-04jointstatement.php3>
- <sup>27</sup> Cronin, J., Kingsbury, G.G., McCall, M.S., & Bowe, B. (2005, November). The impact of the no child left behind act on student achievement and growth: 2005 edition. *Research Brief*. Lake Oswego, OR: Northwest Evaluation Association.
- <sup>28</sup> Ahlgren, A. (1999, July 14). *American Association for the Advancement of Science memorandum* [Commentary]. American Association for the Advancement of Science. Retrieved March 13, 2006, from <http://www.project2061.org/publications/articles/ahlgren4.htm>
- American Education Research Association. (2003). Standards and tests: keeping them aligned. *Research Points*. 1(1), Washington: Author
- Darling-Hammond, L. (2003, February 16). Standards and assessments: where we are and what we need. *Teachers College Record*. Retrieved December 27, 2005, from <http://www.tcrecord.org/Content.asp?ContentID=11109>
- <sup>29</sup> Herman, J.L., Webb, N.M., Zuniga, S.A. (2005). *Measurement issues in the alignment of standards and assessments*. Los Angeles: National Center for Research on Education, Standards and Student Testing (CRESST), CSE report 653.
- <sup>30</sup> Linn, R. L. (2005) *Issues in the design of accountability systems* [CSE technical report 650]. Los Angeles: National Center for Research on Education, Standards and Student Testing (CRESST).

See also,

Rothman, R., Slattery, J.B., Vranek, J.L. and Resnick, L. (2002). *Benchmarking and alignment of standards and testing* [CSE technical report 566]. Los Angeles: National Center for Research on Education, Standards and Student Testing (CRESST).

<sup>31</sup> Pearson, P.D. (2004). The reading wars. *Educational Policy*, 18(1), 216-252.

Also see *Phi Delta Kappan*, June 2002, for diverse articles on this topic.

<sup>32</sup> Kane, T.J. & Staiger, D.O. (2002). The promise and pit falls of using imprecise school accountability measures. *Journal of Economic Perspectives*, 16(4), 91-114.

<sup>33</sup> Herschberg, T. (2005). Value-added assessment and systemic reform: A response to the challenge of human capital development. *Phi Delta Kappan*, 87(4), 276-283.

<sup>34</sup> Kober, N. (2002, October). *What tests can and cannot tell us*. *Testtalk for Leaders*. Washington: Center on Education Policy.

Tucker, A. (2003). *Problems with standards-based mathematics tests*. Retrieved February 9, 2006, from <http://www.ams.sunysb.edu/~tucker/StandardsProb.pdf>

See also,

Tucker, A. (2004, August 11). In testing, how reliable are year-to-year comparisons? *EdWeek.org*. Retrieved May 22, 2005, from <http://www.edweek.org/ew/articles/2004/08/11/44tucker.h23.html?querystring=Tucker>

<sup>35</sup> Tucker, A. (2003). Problems with standards-based mathematics tests. Retrieved February 9, 2006, from <http://www.ams.sunysb.edu/~tucker/StandardsProb.pdf>

- <sup>36</sup> Tucker, A. (2004, August 11). In testing, how reliable are year-to-year comparisons? *EdWeek.org*. Retrieved May 22, 2005, from <http://www.edweek.org/ew/articles/2004/08/11/44tucker.h23.html?querystring=Tucker>
- Tucker, A. (2003). Untitled manuscript. Retrieved February 9, 2006 from <http://www.ams.sunysb.edu/~tucker/StandardsProb.pdf>
- <sup>37</sup> Harris, D. & Herrington, C. (2004, March). *Accountability, standards, and the growing achievement gap: Lessons from the past half-century*. Paper presented at the American Education Finance Association Meeting, Salt Lake City.
- <sup>38</sup> Nichols, S.L., Glass, G.V, & Berliner, D.C. (2005, September). *High-stakes testing and student achievement: Problems for the No Child Left Behind Act*. Tempe: Education Policy Research Unit, Education Policy Studies Laboratory, Arizona State University. Retrieved on January 15, 2006, from <http://www.asu.edu/educ/eps/EPRU/documents/EPSL-0509-105-EPRU.pdf>
- <sup>39</sup> Ritter, G.W., et al. (2004, October 30). *Keeping score: What do we know about the impact of statewide accountability programs on student achievement?* Paper presented at the 2004 Annual meeting of the Association for Public Policy Analysis and Management, Atlanta, GA
- <sup>40</sup> Starr, Alexandra (2005, Aug. 8). They're not stupid—they're lazy. *Slate*. Retrieved May 11, 2006, from <http://www.slate.com/id/2124163/>
- <sup>41</sup> Koretz, D. (2003, April 22). Attempting to discern the effects the NCLB accountability provisions on learning. Presidential invited symposium at the American Educational Research Association, Chicago.
- <sup>42</sup> Center on Education Policy (2005, August). *Do you know the good news about American education?* Washington: Author.
- <sup>43</sup> Center on Education Policy (2005, August). *Do you know the good news about American education?* Washington: Author.
- <sup>44</sup> Education Trust (2005, October 19). *Closing the achievement gap: 2005 NAEP reading and math results show some gains but slowing progress* [Press release]. Retrieved on March 29, 2006, from <http://www2.edtrust.org/EdTrust/Press+Room/NAEP+2005.htm>
- Council of the Great City Schools (2004). *Beating the odds IV: A city by city analysis of student performance and achievement gaps on state assessments*. Washington, DC: Author.
- Education Trust (2006). *Primary progress, secondary challenge: A state-by-state look at student achievement patterns*. Washington, DC: Author. Retrieved on March 12, 2006, from <http://www2.edtrust.org/EdTrust/Press+room/PPSCAdvisory.htm>
- <sup>45</sup> Olson, L. (2005, October 26). NAEP gains are elusive in key areas. *Education Week*, 25(9), pp 1, 22.
- <sup>46</sup> Koretz, D. (2003). *Attempting to discern the effects the NCLB accountability provisions on learning*. Presidential invited symposium at the American Educational Research Association's annual meeting. Chicago.
- <sup>47</sup> Center on Education Policy (2005, August). *Do you know the good news about American education?* Washington, DC: Author.
- Olson, L. (2005, October 26). NAEP gains are Elusive in Key Areas. *Education Week*, 25(9), pp 1, 22.
- Center on Education Policy (2006, January). *The good news behind average NAEP scores*. Washington, DC: Author.
- <sup>48</sup> Linn, R. (2000). Assessments and accountability. *Educational Researcher*, 29(2). 4-16.
- <sup>49</sup> For example, see:

McDade, P.J. (2006, January). *The state of high school education in Wisconsin: A tale of two Wisconsins*. Milwaukee, WI: Wisconsin Policy Research Institute. Retrieved February 1, 2006, from [http://www.wpri.org/Reports/Volume\\_19/Vol19no1.pdf](http://www.wpri.org/Reports/Volume_19/Vol19no1.pdf)

Also see

Sunderman, G.L. & Kim, J. (2004, February) *Large mandates and limited resources: state response to the No Child Left Behind Act and implications for accountability*. The Civil Rights Project. Cambridge: Harvard University.

<sup>50</sup> Center on Education Policy (2006, March). *From the capital to the classroom; Year 4 of the No Child Left Behind Act*. Washington DC: Author.

<sup>51</sup> National Conference of State Legislatures (2005, February). *Task force on No Child Left Behind: Final report* (ISBN 1-58024-383-5). Denver, CO: Author, p.18.

Also see

Center on Education Policy (2006, March). *From the capital to the classroom; Year 4 of the No Child Left Behind Act*. Washington DC: Author.

<sup>52</sup> Kim, J.S. & Sunderman, G.L. (2005, November). Measuring academic proficiency under the No Child Left Behind Act: Implications for educational equity. *Educational Researcher*, 34(8), 3-13.

Linn, R.L. (2003). Accountability: Responsibility and reasonable expectations. *Educational Researcher*, 32(7), 3-13.

<sup>53</sup> Sunderman, G.L. (2006, February). *The unraveling of no child left behind: How negotiated changes transform the law*. Cambridge MA: The Civil Rights Project, Harvard University.

<sup>54</sup> Center on Education Policy (2005, March). *From the capital to the classroom: Year 3 of the No Child Left Behind Act*. Washington, DC: Author.

<sup>55</sup> Gardner, H. (2005, September 14). Beyond the herd mentality: The minds that we truly need in the future. *Education Week*, 25(3). p.44.

Parker, W.C. (2005, January), Teaching against idiocy. *Phi Delta Kappan*, 86(5), 344-351.

Darling-Hammond, L. (2003, February 16). *Standards and assessment: Where we are and what we need*. New York: Teachers College Record Online. Retrieved December 27, 2005, from <http://www.tcrecord.org/Content.asp?ContentID=11109>

<sup>56</sup> Darling-Hammond, L (2003, February 16). *Standards and assessment: Where we are and what we need*. New York: Teachers College Record Online. Retrieved December 27, 2005, from <http://www.tcrecord.org/Content.asp?ContentID=11109>

<sup>57</sup> Center on Education Policy (2005, July). *NCLB: Narrowing the curriculum? NCLB policy brief 3*. Washington, DC: Author.

<sup>58</sup> Center on Education Policy (2006, March). *From the capital to the classroom: Year 4 of the No Child Left Behind Act*. Washington, DC: Author.

<sup>59</sup> Policy Research Shop (2005). *Curriculum shifts in New Hampshire and Vermont public schools: A survey of school superintendents*. Power point presentation to Vermont Joint NCLB Oversight Committee. Hanover, NH: Nelson A. Rockefeller Center, Dartmouth College.

<sup>60</sup> Hargreaves, A. (2003). *Teaching in the Knowledge Society: Education in the Age of Insecurity*. New York: Teachers College Press.

<sup>61</sup> Haycock, K. (2005, September 29) *Closing the achievement gap in America's public schools: the No Child Left Behind Act*. Testimony before the House Committee on Education and the Workforce. Washington, DC: Education Trust.



- <sup>62</sup> National Conference of State Legislatures (2005, February). *Task force on No Child Left Behind: Final report* (ISBN 1-58024-383-5). Denver, CO: Author.
- <sup>63</sup> Center on Education Policy (2006, March). *From the capital to the classroom: Year 4 of the No Child Left Behind Act*. Washington, DC: Author.
- <sup>64</sup> Education Trust (2005, Winter). *The funding gap 2005: Low-income and minority students shortchanged by most states*. Washington, DC: Author.
- <sup>65</sup> Paige, R. (2003, October 29). It's not about the money [Guest opinion article]. *The Wall Street Journal*.
- <sup>66</sup> Paige, R. (2004, May 25). *New GAO report finds that no child left behind is not an unfunded mandate*. [Press release]. Washington DC: U. S. Department of Education.
- <sup>67</sup> Boehner, J. (2004, February 12). *New national analysis shows states profiting financially from the no child left behind act* [Press Release]. Washington DC: U.S. House of Representatives.
- <sup>68</sup> Mathis, W. (2005). The cost of implementing the federal No Child Left Behind Act: Different assumptions, different answers. *Peabody Journal of Education*. 80(2), pp 91-120.
- <sup>69</sup> American Federation of Teachers (2004, January 12) *New Bush FY2005 budget for No Child Left Behind comes up short* [Press release]. Washington, DC: Author.
- National Education Association (2004). *No child left behind? The funding gap in ESEA and other federal programs*. Washington, DC: Author.
- <sup>70</sup> Palaich, R. (Personal email communication regarding the CCSSO NCLB administrative cost studies, February 28, 2006).
- Also see,
- Mathis, W. (2005). The cost of implementing the No Child Left Behind Act: Different assumptions, different answers. *Peabody Journal of Education*, 80(2), 90-118.
- <sup>71</sup> Davis, M.R. (2006, February 15). President's budget would cut education spending. *Education Week*, 25(23), pp.1, 24.
- <sup>72</sup> Peyser, J. & Costrell R. (2004, Spring). Exploring the Costs of Accountability. *EducationNext*. Retrieved June 11, 2004, from <http://www.educationnext.org/20042/22.html>
- <sup>73</sup> Rothstein, R. (2004). *Class and schools: using social, economic and educational reform to close the achievement gap*. New York: Economic Policy Institute/Teachers' College.
- Berliner, D. (2005, May 25). *Our impoverished view of educational reform*. Presidential address at the American Educational Research Association's Annual Meeting, Montreal.
- See also:
- Barton, P.E. (2003, October). *Parsing the achievement gap: Baselines for tracking progress*. Princeton, NJ: Education Testing Service.
- For more on this topic see papers from *The Social Costs of Inadequate Education* (Fall, 2005), a symposium at Columbia University. Available at <http://www.tc.edu/symposium/>
- <sup>74</sup> Mathis, W.J.(2005). The cost of implementing the No Child Left Behind Act: Different assumptions, different answers. *Peabody Journal of Education*, 80(2), 90-118.
- <sup>75</sup> Hanushek, E.A. (2005). *The alchemy of costing out an adequate education*. Cambridge, MA: Kennedy School of Governance conference on "Adequacy Lawsuits: Their Growing Impact on American Education."
- <sup>76</sup> Taylor, L.L., Gronberg, T.J., Jansen, D. W., & Booker, K. (2004). *School outcomes and school costs: The cost function approach*. College Station, TX: Texas A&M University.



- <sup>77</sup> Education Trust. (2005, Winter). *The funding gap 2005: Low-income and minority students shortchanged by most states*. Washington, DC: Author.
- <sup>78</sup> Wiley, E., Mathis, W., & Garcia, D. (2005, May). *The impact of the adequate yearly progress requirement of the federal "No Child Left Behind" act on schools in the Great Lakes region*. Tempe, AZ: Education Policy Research Unit, Education Policy Studies Laboratory, Arizona State University. Retrieved from <http://www.asu.edu/educ/eps/EPRU/documents/ERSL-0509-109-EPRU-app.pdf>
- <sup>79</sup> Lee, J. (2003, April 22). *How Adequate Yearly Progress works: Simulation analyses of Maine and Kentucky school AYP data*. Paper presented at AERA annual meeting, Chicago.
- <sup>80</sup> Center on Education Policy (2005, March). *From the capital to the classroom: Year 3 of the No Child Left Behind Act*. Washington, DC: Author
- <sup>81</sup> Olson, L. (2004, September 8). Data show schools making progress on federal goals. *Education Week*. Retrieved May 4, 2005, from <http://www.edweek.org/ew/articles/2004/09/08/02ayp.h24.html>
- <sup>82</sup> Feller, Ben (2006, May 10). GOP joins inquiry about uncounted kids. *Associated Press*. Retrieved May 18, 2006 at [http://hosted.ap.org/dynamic/stories/N/NO\\_CHILD\\_LOOPHOLE?SITE=AP&SECTION](http://hosted.ap.org/dynamic/stories/N/NO_CHILD_LOOPHOLE?SITE=AP&SECTION)
- <sup>83</sup> Feller, Ben (2006, May 10). More schools face drastic action under No Child Left Behind Law. *AP/Cleveland Plain Dealer*. Retrieved May 10, 2006, from <http://www.cleveland.com/news/plaindealer/index.ssf?/base/news/114726070060330.xml&coll=2>
- <sup>84</sup> See, for example:
- Haycock, K. (2005, September 29). *Closing the achievement gap in America's public schools: The No Child Left Behind Act*. Testimony before the House Committee on Education and the Workforce. Washington: Education Trust.
- Butler, C.K. (2005, October 3) Success in the City. *U.S. News and World Report*. Retrieved March 13, 2006, from <http://www.usnews.com/usnews/edu/articles/051003/3district.htm>
- <sup>85</sup> Haney, W. (2002). Lake Woebeguaranteed: Misuse of test scores in Massachusetts; Part I. *Education Policy Analysis Archives*, 10(6). Retrieved February 10, 2006, from <http://epaa.asu.edu/epaa/v10n24>
- <sup>86</sup> Harris, D.N. (2006, March) *Ending the blame game on educational inequity*. Tempe, AZ: Educational Policy Research Unit, Education Policy Studies Laboratory, Arizona State University. Retrieved April 16, 2006, from <http://www.asu.edu/educ/eps/EPRU/documents/EPSL-0603-120-EPRU.pdf>
- <sup>87</sup> Sunderman, G.L. (2006, February). *The unraveling of No Child Left Behind: How negotiated changes transform the law*. Cambridge, MA: The Civil Rights Project, Harvard University.
- <sup>88</sup> Hanushek, E. A. (2005). *The alchemy of costing out an adequate education*. Cambridge MA: Kennedy School of Governance conference on "Adequacy Lawsuits: Their Growing Impact on American Education."
- <sup>89</sup> Taylor, L.L., Gronberg, T.J., Jansen, D.W., & Booker, K. (2004). *School outcomes and school costs: The cost function approach*. College Station, TX: Texas A&M University.
- <sup>90</sup> Center on Education Policy (2005, March). *From the capital to the classroom: Year 3 of the No Child Left Behind Act*. Washington, DC: Author.
- Haycock, K. (2005, September 29) *Closing the achievement gap in America's public schools: The No Child Left Behind Act*. Testimony before the House Committee on Education and the Workforce. Washington, DC: Education Trust.
- <sup>91</sup> Barton, P.E. (2003, October). *Parsing the achievement gap: Baselines for tracking progress*. Princeton, NJ: Education Testing Service.

For more on this topic see papers from *The Social Costs of Inadequate Education* (Fall, 2005), a symposium at Columbia University. Available at <http://www.tc.edu/symposium/>

<sup>92</sup> Hargreaves, A. (2003). *Teaching in the knowledge society: Education in the age of insecurity*. New York: Teachers' College Press.

Kohn, A. (2004, April). Test today, privatize tomorrow. *Phi Delta Kappan*, 85(8), 568-577. Retrieved March 13, 2006, from <http://www.alfiekohn.org/teaching/testtoday.htm>

---

The Arizona State University Education Policy Research Unit  
Editorial Review Board reviewed this document.

For information on the board and its members, visit

<http://www.asu.edu/educ/epsI/ReviewBoard/reviewboard.htm>