

Testimony before the No Child Left Behind Commission

Hearing on the Quality and Consistency of State Standards

Cambridge, Massachusetts

August 31, 2006

Chester E. Finn, Jr.

President, Thomas B. Fordham Foundation

Senior Fellow, Hoover Institution, Stanford University

Thank you for inviting me to speak with you today about the No Child Left Behind Act (NCLB) and the quality and consistency of state standards. The commission staff encouraged me to tackle four questions in this area, which I'm pleased to do.

Quality of State Standards

First, the "relative quality and rigor of state standards in reading/language arts and math that are being used as the basis for state accountability systems under NCLB."

As you know, standards-based reform is one of the two driving engines of education improvement in the United States and has been at least since 1989. (The other engine is school choice in its innumerable varieties.) Though many states commenced this process on their own, federal encouragement has caused them all to do so, beginning with the Improving America's Schools Act and the Goals 2000 Act, both passed in 1994, then NCLB in 2001. Over the past decade, 49 states and the District of Columbia have created, replaced, substantially revised, or augmented their English and math standards. NCLB, of course, raised the stakes attached to those standards. States, districts and schools are now judged by how well they educate their students in relation to those standards. Moreover, billions of dollars in federal aid now hinge on whether states hold their schools and districts to account for student learning as defined in those standards and measured on assessments that are supposed to be aligned with those standards.

Thus a state's academic standards bear far greater weight than ever before. They now constitute the foundation for a complex, high-stakes accountability system, the first and arguably most important leg in the tripod of "standards," "tests," and "accountability." And that's only the beginning. In well functioning education systems, these standards are also the underpinning for teacher preparation and professional development, textbook selection, and much more. If this foundation is sturdy, such reforms may succeed in boosting student achievement; if it's weak, uneven, or cracked, reforms erected atop it will be shaky and, in the end, may collapse, proving to be worse than none at all.

Let me be frank: with a handful of laudable exceptions, the academic standards in use in most states today range from mediocre to dreadful. My foundation has undertaken systematic reviews of state standards in various subjects. Here's an overview of the results for math and reading. (We can also supply current information on the quality of state standards in science, U.S. history, and world history.)

In our 2005 report, *The State of State English Standards*, written by Massachusetts's own Sandra Stotsky, just five states received A grades, while eight received marks of D or F. Nearly half got Cs. Yes, this was better than in Stotsky's previous review of state English/language arts standards, conducted in 2000. In fact, English/language arts (and reading) is the subject in which states have most improved their standards. But many of them remain lamentable.

2005 State English Standards—ranked by grade

Table 1: 2005 Final Grades

by state rank

State	GPA	Grade	Rank
Massachusetts	3.91	A	1
California	3.68	A	2
Alabama	3.64	A	3
Louisiana	3.59	A	4
Indiana	3.50	A	5
South Dakota	3.36	B	6
Georgia	3.27	B	7
Virginia	3.23	B	8
Minnesota	3.14	B	9
Texas	3.14	B	9
Illinois	3.09	B	11
North Carolina	3.05	B	12
Arizona	2.91	B	13
New Hampshire	2.91	B	13
South Carolina	2.91	B	13
Idaho	2.82	B	16
Mississippi	2.82	B	16
New York	2.82	B	16
Nevada	2.77	B	19
Oregon	2.77	B	19
North Dakota	2.68	C	21
Pennsylvania	2.68	C	21
West Virginia	2.68	C	21
Oklahoma	2.64	C	24
Ohio	2.55	C	25
Maryland	2.45	C	26
Maine	2.45	C	26
National Average: 2.41			
Vermont	2.41	C	28
Missouri	2.36	C	29
Arkansas	2.32	C	30
Washington, DC	2.32	C	30
Nebraska	2.27	C	32
Rhode Island	2.09	C	33
Delaware	2.05	C	34
Florida	2.05	C	34
Wisconsin	2.00	C	36
New Jersey	1.95	C	37
Hawaii	1.91	C	38
Kansas	1.91	C	38
Colorado	1.82	C	40
Kentucky	1.77	C	41
Utah	1.73	C	42
Alaska	1.68	D	43
New Mexico	1.59	D	44
Tennessee	1.45	D	45
Michigan	1.41	D	46
Wyoming	1.27	F	47
Washington	1.23	F	48
Connecticut	1.09	F	49
Montana	0.82	F	50
Iowa	*	*	*

* Indicates the state had no standards at the time of this review.

What was Dr. Stotsky looking for? Standards were evaluated in a number of areas, but her criteria fall into six major groups: 1) Quality of their content (Do English standards provide solid literature content and identify major authors and works that students should read?); 2) Are standards measurable and teachable (Do they ask students to perform specific tasks or understand specific knowledge or do deal in platitudinous language and intangible goals?); 3) Is American literature specifically identified or does it go unmentioned?; 4) Are standards presented grade-by-grade (as they should be) or grouped into unmanageably wide “grade spans”?; 5) Do standards distinguish higher-level concepts from lower-level skills?; and 6) Do they avoid imposing moral or sociological dogma in the classroom? (The full list of her criteria is included as an attachment.)

Unfortunately, most state standards didn’t make the grade. At least not a very high grade.

There are, to be sure, some bright spots on this horizon. Standards for early reading instruction have improved since enactment of NCLB. Looking across all the states, we found substantial gains, especially in grades 3-8 reading standards, which bear the heaviest weight under NCLB. On a four-point scale, the average state grade in this area rose from 1.98 in 2000 to 2.41 in 2005. Most states have also heeded the emerging research consensus on early reading instruction and are incorporating the recommendations of the National Reading Panel into their standards, including systematic phonics instruction. Overall, they do a better job of addressing listening, reading, and writing skills and strategies than five years earlier.

On the other hand, literature remains sorely neglected—more so, in fact, than before, particularly at the high-school level. This is now the great weakness in state English standards, perhaps because NCLB focuses on grades 3-8. Uncorrected, it portends a generation of Americans who may know how to read but, by the end of high school, cannot be assumed to have read much that’s worthwhile, let alone acquired a suitable grounding in the great works of our shared culture.

As for math, Fordham’s 2005 appraisal, *The State of State Math Standards*, was even more critical. The average state grade was a D. Only 6 states received A’s or B’s.

2005 State of State Math Standards—Ranked by grade

Fig. 2: 2005 Results, ranked

	Score	Grade
California	3.89	A
Indiana	3.82	A
Massachusetts	3.30	A
Alabama	2.97	B
New Mexico	2.67	B
Georgia	2.53	B
West Virginia	2.35	C
New York	2.08	C
Michigan	2.00	C
Arizona	2.00	C
Oklahoma	1.97	C
Virginia	1.97	C
North Carolina	1.82	C
South Dakota	1.80	C
Texas	1.80	C
Illinois	1.80	C
Kentucky	1.80	C
North Dakota	1.80	C
Louisiana	1.78	C
Maryland	1.77	C
Nevada	1.77	C
Tennessee	1.70	D
Minnesota	1.67	D
Mississippi	1.67	D
National Average	1.59	D
Wisconsin	1.50	D
Ohio	1.43	D
Nebraska	1.42	D
Colorado	1.37	D
DC	1.37	D
Maine	1.35	D
Oregon	1.35	D
Alaska	1.32	D
South Carolina	1.32	D
Pennsylvania	1.28	D
Vermont	1.20	D
New Jersey	1.15	D
Utah	1.13	D
Idaho	1.10	D
Montana	1.00	D
Wyoming	0.98	F
Florida	0.93	F
Kansas	0.83	F
Arkansas	0.72	F
New Hampshire	0.70	F
Rhode Island	0.67	F
Missouri	0.57	F
Washington	0.57	F
Delaware	0.54	F
Connecticut	0.47	F
Hawaii	0.43	F
Iowa	-	-

What were our math reviewers—a group of respected mathematicians, led by Professor David Klein of California State University-Northridge—looking for? First, and most important, they wanted to see rigorous, appropriate, and accurate math content. For example, states should require children to know the basic number facts and have facility with the standard algorithms of arithmetic. They also sought standards that are clear, that expect students to demonstrate strong mathematical reasoning, and to be free of negative qualities like relying overmuch on manipulatives and calculators. But content is king. After all, Klein explained, if solid content isn't there (or is wrong), such factors as "clarity of expression" cannot compensate. Such standards resemble clearly written recipes that use the wrong ingredients or combine them in the wrong proportions. (The full list of their criteria is included as an attachment.)

Math standards have gotten notably worse since the prior Fordham review. One of the most debilitating trends is their excessive emphasis on calculators. In Idaho, children use them beginning in kindergarten! Most standards documents expect students to use them starting in the elementary grades. Calculators enable students to do arithmetic without thinking about the numbers involved in a calculation. For this reason, it makes sense to use them in, say, a high school science class. But for elementary students, the main goal of math education is to get them to think about numbers and to learn arithmetic. Excessive and premature reliance on calculators defeats that purpose.

Besides applying the criteria and rendering judgments on the standards, Klein and his team identified a set of widespread failings that weaken math standards in many states. States do not require the memorization of basic number facts, they pay too little attention to teaching fractions, they obsess over having students identify patterns, and they do not develop strategies for solving word problems in an effective way. The authors also trace the source of much of this weakness to states' unfortunate embrace of the advice of the National Council of Teachers of Mathematics (NCTM), particularly that organization's wrongheaded 1989 standards. (A later NCTM publication made

partial amends, but these came too late for the standards—and children—of many states.)

In sum, too many states have weak standards in English/ Language Arts and worse ones in math. Has No Child Left Behind had any impact on this situation? On the positive side, grades for state English standards have improved somewhat since our last review in 2000, reflecting their stronger focus on scientifically-based reading in the early grades. For this development, NCLB and its Reading First program can justifiably take some credit. Math standards, however, have gotten worse.

	2000 Average	2006 Average	2000 Honors (# of states)	2006 Honors (# of states)
English	C-	C+	19	20
Mathematics	C	D+	18	6

Of course, formal, written academic standards tell only part of the story. States may have standards that look great on paper yet still game the NCLB system in other ways. For instance, they may decide to test students only on the easier parts of their standards. Or they may alter the mix of test questions, de-emphasizing the tougher ones. And, of course, states can set the cut scores on their state tests at low levels. That way, almost every student passes—even those who haven’t come close to mastering the state’s lofty standards. These kinds of finaglings are hard to detect and can generally be done in secret.

Good standards aren’t the solution to every education problem but they do matter. As detailed in a brand-new Fordham report, *The State of State Standards 2006*, only seven states made statistically significant progress from 1998 to 2005 in the percentage of students reaching proficiency in fourth-grade reading, and just six states made such

gains among poor or minority students. All but one of these states had received at least a “C” from Fordham for their English/Language Arts standards (see table below). That’s not iron-clad proof that good standards boost achievement, but it certainly suggests that bad standards make it less likely.

**States Making Statistically Significant Gains on the
4th Grade Reading NAEP, 1998-2005**

	Gains for All Students	Gains for Low Income or Minority Students	Fordham Grade for English/ Language Arts Standards
Arkansas	X	X	C
California		X	A
Delaware	X	X	C
Florida	X	X	C
Hawaii	X		C
Massachusetts	X		A
New York		X	B
Utah	X		C
Virginia	X		B
Wyoming		X	F

How NCLB Could Address Faulty State Standards

The commission staff asked me to consider “what role NCLB should play” in addressing the “differences in focus and rigor among state standards.” That’s a difficult question, and as I explain below, the surest response is to move toward national standards and tests. But let me suggest some lesser adjustments to NCLB that might help.

One is to have the Federal government carry out alignment studies to see how states’ standards match up to each another and to the content “frameworks” that guide NAEP.

At the very least, NAEP's role as a benchmark exam could be further strengthened. A recent Fordham analysis demonstrates how this might work. We looked at trends on state tests over time and compared them to trends of state performance on NAEP. This provides some indication of whether state standards and tests are getting easier or harder over time. After all, if students show big gains on the state test, at least some of those gains ought to show up on an external benchmark exam like NAEP (although one would expect stronger performance on the state assessment since it is a high-stakes exam aligned to state standards and curricula, and NAEP is not). When Fordham performed such an analysis last year, we found that, from 2003 to 2005, at least 20 states (of the 30-odd for which we could gather data) posted gains on their own 8th-grade reading exams, yet *none* of these showed progress at the "proficient" level on NAEP. Only three showed progress at even the "basic" level. Alabama, for example, demonstrated on its state test an 11-point increase from 2003 to 2005 in eighth graders reading at or above proficient. Yet over the same time period NAEP showed no change in proficient eighth-grade Alabama readers. NAEP also showed that the percentage of eighth graders in Alabama reading at a basic level *dropped* two points. While there could be other explanations, one must suspect that some state tests are getting easier over time.

Gains on State Tests vs. Gains on NAEP: 8th Grade Reading: 2003-2005

State	Change on State Test : Percentage of Students at or Above Proficient	Change on NAEP: Percentage of Students at or Above Basic	Change on NAEP: Percentage of Students at or Above Proficient
Alabama	11	-2	0
California	9	-1	-1
Idaho	9	0	0
Arizona	8	-1	-2
Delaware	8	3	-1
Tennessee	8	2	0
Maryland	6	-2	-1
Virginia	6	-1	0
Kentucky*	5	-3	-3
Indiana	3	-4	-5
Iowa	3	0	-2
New York	3	0	-2
North Dakota	3	2	-1
Oregon	3	-1	0
Georgia	2	-2	-1
North Carolina	2	-3	-2
Oklahoma	2	-2	-5
Missouri	1	-3	-3
South Dakota	1	0	-4
Colorado	0	-3	-4
Mississippi*	0	-5	-4
Wisconsin	0	0	-2
Wyoming	0	2	2
Hawaii	-1	-3	-4
Maine	-1	2	0
Connecticut	-2	-3	-4
Florida	-5	-2	-2
Texas	-5	-2	0
Montana	-6	0	0
Alaska	N/A	3	0
Arkansas	N/A	-1	-1
Illinois	N/A	-2	-4
Kansas	N/A	1	0
Louisiana	N/A	0	-2
Massachusetts	N/A	2	1
Michigan*	N/A	-2	-4
Minnesota*	N/A	2	0
Nebraska	N/A	3	0
Nevada	N/A	0	1
New Hampshire	N/A	-1	-2
New Jersey	N/A	1	1
New Mexico	N/A	0	-1
Ohio	N/A	0	2
Pennsylvania	N/A	1	4
Rhode Island	N/A	0	-1
Utah	N/A	-3	-3
Vermont	N/A	-2	-2
Washington*	N/A	-1	1
West Virginia	N/A	-5	-3

* State did not offer reading test at grade 4. *Education Week* accepted test results from the next closest grade level.

Again, if Washington were to perform this type of analysis and publicize it widely, it might lead some states to change their behavior and "race to the top."

Learning from NAEP

The commission staff asked me to address “what NAEP results reveal about the relative quality, consistency and rigor of state standards.” I already explained how looking at NAEP trends versus state test score trends can illuminate whether state tests are getting easier or harder. But NAEP, while not a perfect metric, is also useful for comparing the absolute rigor of state tests and their definitions of “proficiency.”

Everyone knows that there’s a huge gap between the percentage of students scoring “proficient” on state tests and the percentage scoring at that level on NAEP. For instance, in 2005, Tennessee reported that 87 percent of its fourth-graders are proficient in reading while NAEP reports that only 27 percent of Tennessee children meet that level.

Paul Peterson and Frederick Hess have performed a helpful analysis (in *Education Next*) that indicates which states have the loftiest or lowest definitions of “proficiency” by comparing state assessment results to NAEP results. The grades reported here (see next page) are based on a comparison of state and NAEP proficiency scores in 2005, with changes calculated relative to 2003. A high grade on this chart indicates that the percentage of a state’s students reaching proficiency on its own exam is similar to the percentage of the states’ students who are proficient on NAEP. In other words, the state’s definition of “proficiency” is about as rigorous as NAEP’s definition. States that do poorly on this chart set their own definition of proficiency perilously low—and artificially inflate the percentage of their students deemed proficient.

Strength of State Proficiency Standards, 2005

	4th grade		8th grade		OVERALL GRADE	
	Math	Reading	Math	Reading	2005	2003
Massachusetts	A	A	A		A	A
Maine	A	A	A	A	A	A
South Carolina	A	A	A	A	A	A
District of Columbia	A	A	B+	A	A	
Wyoming	A	A	B	A	A	A
Missouri	B		A		A	A
Hawaii	A	C+	A	B	B+	B
California	C+	B		B	B-	B
Arkansas	B	B+	B	C	B-	C+
Rhode Island	C+	C	B-	B+	B-	B-
New Mexico	B-	C+	B	C	C+	
Kentucky		C	B-		C+	B-
Montana	B-	C	C	C+	C	C-
Florida	C	C	C-	B	C	C
Washington	B-	C-			C	C+
Nevada			C	C+	C	
New York	D-	C	C	B+	C	C
Oregon			C	C	C	
Pennsylvania			C-	C+	C	C
Ohio	C+	C	C	D+	C	C+
New Jersey	C-	C-	C	C	C	C
North Dakota	C-	C	C	C	C	C
Illinois			C	C-	C	C
Louisiana	C-	C-	C-	C	C	C-
Connecticut	C-	B-	D	C-	C	C-
Maryland	C-	D+	C	C	C	C+
Delaware			C	D	C-	C
Indiana	C-	C	D	C-	C-	C-
Michigan	C	D+	C-		C-	C
Kansas	C-			D+	C-	C-
Wisconsin	C	D+	D+	D	C-	D
Iowa	D+	C-	D	C	D+	
Arizona	D	C-	D+	D+	D+	B-
South Dakota	D+	D	C-	D+	D+	C-
Utah	C-	C-	D	D	D+	
Alaska	C-	D	C-	D-	D+	D+
Texas	D+	D+	C	F	D+	F
Virginia			D-	C-	D+	D+
Idaho	D-	D	D+	D	D	D+
Colorado	D-	D+	D	D-	D	D
Nebraska	D-	D	D-	D-	D-	
Alabama	D-	F	D-	D	D-	
Mississippi	F	F	D+	C-	D-	D-
Georgia	D	F	D-	F	D-	D-
West Virginia	D-	D-	F	F	D-	
North Carolina	D-	D	F	F	F	D-
Oklahoma	F	F	F	F	F	F
Tennessee	F	F	F	F	F	F

Some states are obviously setting the bar higher than others, although it's not possible, just by looking at NAEP results, to know much about the actual content of state standards. For example, Virginia's "Standards of Learning" in most subjects are among the nation's best according to our reviewers. Yet judging by the percentage of students "proficient" on the Virginia test versus the percentage proficient on NAEP, it has one of the lowest "cut scores" for proficiency. Conversely, Wyoming has terrible content standards, but its test (based on who knows what) is pegged at a relatively high level. So looking at NAEP results is not enough.

How NCLB Can Lead to High Expectations for All

Finally, the commission staff requested my "views on the most effective way(s) to assure that NCLB results in high expectations for the academic performance of all children."

I start here with the conclusion that NCLB's architect's made one fundamental mistake. Rather than setting a common standard for school performance across the land and then encouraging states, districts, and schools to meet that standard as they judge best, they allowed states to define "proficiency" in reading and math as they saw fit. Instead of regulating ends, in other words, Washington has once again found itself regulating means, prescribing a hundred different aspects of what states and districts should do when, by their own lights, schools don't do a satisfactory job. That means lots and lots of regulation on the one hand and, on the other, heavy pressure on states to define "proficiency" downward and make Swiss cheese out of NCLB's accountability provisions. Already many states, in order to explain the discrepancy between their passing rates on state tests and their students' performance on NAEP, are claiming that observers should equate state "proficiency" with NAEP's "basic" level. In other words, they are satisfied to get their students to "basic," proficiency be damned. A system that allows such cheese-paring and redefining puts the entire standards-based-reform enterprise in peril.

The surest way to end this gamesmanship—and keep Washington from playing a cat-and-mouse game with recalcitrant states—is to move to a system of national standards and tests, a system that simultaneously frees states, districts, and schools to get to those standards as they see fit.

Such a system could be implemented in many ways. We've identified four of them, detailed in the attached report, *To Dream the Impossible Dream: Four Approaches to National Standards and Tests for America's Schools*. No doubt there are more waiting to be devised. Here are our four:

1. **The Whole Enchilada.** This is the most direct and aggressive approach. Uncle Sam would create and enforce national standards and assessments, replacing the fifty state-level sets of standards and tests we have now. The United States would move to a national accountability system for K-12 education.
2. **If You Build It, They Will Come.** This is a voluntary version of the first model. Washington would develop national standards, tests and accountability metrics, and provide incentives to states (e.g. more money, fewer regulations) to opt into such a system. A variant would have a private group frame the standards. But participation would be optional for states, which would remain free to set their own standards if they prefer.
3. **Let's All Hold Hands.** Under this approach, states would be encouraged to join together to develop common standards and tests or, at the least, common test items. Uncle Sam might provide incentives for such collaboration, but that's it.
4. **Sunshine and Shame.** This model, the least ambitious (and already foreshadowed in my earlier remarks), would make state standards and tests more transparent by making them easier to compare to one another and to the

National Assessment of Educational Progress (NAEP).

To me, the most promising approach is a version of the second model, “If You Build It, They Will Come.” We would charge the National Assessment Governing Board (NAGB) with setting standards—in grades 3-12 reading, math, and science, for starters—and developing world-class tests aligned with those standards and the underlying content frameworks. NAGB would build on existing NAEP frameworks and exams to develop a system of annual tests that would assess individual children in these three subjects in this expanded number of grades. (That’s not as easy as it sounds; NAEP is currently a “matrix sample” test—no child actually takes the entire exam and no scores are computed for individual students or schools. Hence this assignment to NAGB is a very large one.) Neither Congress nor the Administration should play a direct role in approving the frameworks or performance standards.

We admit to mild angst about using NAGB and NAEP for this role. On the one hand, they’re working pretty well as a low-stakes external “audit” of state and national performance and their success as an outside auditor could be compromised by the added burdens envisioned here. Moreover, some responsible critics assert that NAEP’s frameworks in certain subjects, especially math, are as problematic as those of many states. We also know that NAGB is not immune to politicians’ demands to see test scores rise. For example, its recent emphasis on the “basic” NAEP level rather than “proficient” causes us concern. As long as NAGB members are appointed by the Administration in power, these risks will remain.

Yet NAGB has many strengths, too. It is a broadly representative and bipartisan body, with all key stakeholders represented. Its processes are relatively open and transparent. It has not been timid about demonstrating its independence both of political masters and of education interest groups. Perhaps most important, it has experience setting standards and developing a national test—one that is highly regarded. While we might prefer California’s or Massachusetts’ standards to the NAEP frameworks, we appreciate

that NAEP is a reasonable representation of a broad consensus about what American students should know and be able to do, and it is certainly true that NAGB's concept of "proficiency" is as rigorous and challenging as the 21st Century demands. Were Congress to design a new standards-setting body and process from scratch, it would probably look a lot like NAGB and result in a test much like NAEP—but would take a long time to create the infrastructures, culture and working relationships that NAGB already has. Hence relying on NAGB and NAEP is the best way to hit the ground running.

Finally, and perhaps counter-intuitively, we see national standards and tests as an opportunity to *rein in* the federal government. For forty years, Washington has sought to improve the nation's schools by regulating what they do. No Child Left Behind has, by requiring testing in only two subjects, exerted definite pressure on schools to restructure their curricula and emphasize math and reading to the detriment of other subjects. To date, scant evidence exists that this strategy works. Common standards and tests could allow Uncle Sam to back away from his top-down, regulatory approach and settled instead for clarifying the objectives to be achieved and measuring (and publicizing) whether states, schools, and students are in fact meeting them. Many think that national standards entail an increased federal role. We see it in precisely the opposite way: that a good set of national standards will lead to a reduced and focused federal role that is also better suited to Washington's limited skill set.

Thanks once again for giving me the opportunity to share my views on this topic.

Attachments

- Criteria for Fordham reviews of state English and math standards
- *To Dream the Impossible Dream: Four Approaches to National Standards and Tests for America's Schools*