REAUTHORIZATION OF THE ELEMENTARY AND SECONDARY EDUCATION ACT: CURRENT AND PROSPECTIVE FLEXIBILITY UNDER NO CHILD LEFT BEHIND

HEARING

BEFORE THE

SUBCOMMITTEE ON EARLY CHILDHOOD, ELEMENTARY AND SECONDARY EDUCATION

COMMITTEE ON EDUCATION AND LABOR

U.S. HOUSE OF REPRESENTATIVES

ONE HUNDRED TENTH CONGRESS

FIRST SESSION

HEARING HELD IN WASHINGTON, DC, JUNE 7, 2007

Serial No. 110–46

Printed for the use of the Committee on Education and Labor

Available on the Internet:
http://www.gpoaccess.gov/congress/house/education/index.html

U.S. GOVERNMENT PRINTING OFFICE

35–664 PDF WASHINGTON : 2008

For sale by the Superintendent of Documents, U.S. Government Printing Office
Internet: bookstore.gpo.gov Phone: toll free (866) 512–1800; DC area (202) 512–1800
Fax: (202) 512–2104 Mail: Stop IDCC, Washington, DC 20402–0001

COMMITTEE ON EDUCATION AND LABOR

GEORGE MILLER, California, *Chairman*

Dale E. Kildee, Michigan, *Vice Chairman*
Donald M. Payne, New Jersey
Robert E. Andrews, New Jersey
Robert C. "Bobby" Scott, Virginia
Lynn C. Woolsey, California
Rubén Hinojosa, Texas
Carolyn McCarthy, New York
John F. Tierney, Massachusetts
Dennis J. Kucinich, Ohio
David Wu, Oregon
Rush D. Holt, New Jersey
Susan A. Davis, California
Danny K. Davis, Illinois
Raúl M. Grijalva, Arizona
Timothy H. Bishop, New York
Linda T. Sánchez, California
John P. Sarbanes, Maryland
Joe Sestak, Pennsylvania
David Loebsack, Iowa
Mazie Hirono, Hawaii
Jason Altmire, Pennsylvania
John A. Yarmuth, Kentucky
Phil Hare, Illinois
Yvette D. Clarke, New York
Joe Courtney, Connecticut
Carol Shea-Porter, New Hampshire

Howard P. "Buck" McKeon, California, *Ranking Minority Member*
Thomas E. Petri, Wisconsin
Peter Hoekstra, Michigan
Michael N. Castle, Delaware
Mark E. Souder, Indiana
Vernon J. Ehlers, Michigan
Judy Biggert, Illinois
Todd Russell Platts, Pennsylvania
Ric Keller, Florida
Joe Wilson, South Carolina
John Kline, Minnesota
Cathy McMorris Rodgers, Washington
Kenny Marchant, Texas
Tom Price, Georgia
Luis G. Fortuño, Puerto Rico
Charles W. Boustany, Jr., Louisiana
Virginia Foxx, North Carolina
John R. "Randy" Kuhl, Jr., New York
Rob Bishop, Utah
David Davis, Tennessee
Timothy Walberg, Michigan
Dean Heller, Nevada

Mark Zuckerman, *Staff Director*
Vic Klatt, *Minority Staff Director*

SUBCOMMITTEE ON EARLY CHILDHOOD, ELEMENTARY AND SECONDARY EDUCATION

DALE E. KILDEE, Michigan, *Chairman*

Robert C. "Bobby" Scott, Virginia
Dennis J. Kucinich, Ohio
Susan A. Davis, California
Danny K. Davis, Illinois
Raúl M. Grijalva, Arizona
Donald M. Payne, New Jersey
Rush D. Holt, New Jersey
Linda T. Sánchez, California
John P. Sarbanes, Maryland
Joe Sestak, Pennsylvania
David Loebsack, Iowa
Mazie Hirono, Hawaii
Phil Hare, Illinois
Lynn C. Woolsey, California
Rubén Hinojosa, Texas

Michael N. Castle, Delaware, *Ranking Minority Member*
Peter Hoekstra, Michigan
Mark E. Souder, Indiana
Vernon J. Ehlers, Michigan
Judy Biggert, Illinois
Luis G. Fortuño, Puerto Rico
Rob Bishop, Utah
Todd Russell Platts, Pennsylvania
Ric Keller, Florida
Joe Wilson, South Carolina
Charles W. Boustany, Jr., Louisiana
John R. "Randy" Kuhl, Jr., New York
Dean Heller, Nevada

C O N T E N T S

	Page
Hearing held on June 7, 2007	1
Statement of Members:	
Castle, Hon. Michael N., Senior Republican Member, Subcommittee on Early Childhood, Elementary and Secondary Education	3
Prepared statement of	5
Hare, Hon. Phil, a Representative in Congress from the State of Illinois, submission for the record:	
Prepared statement of the Coalition Promoting School Success for All Children	20
Kildee, Hon. Dale E., Chairman, Subcommittee on Early Childhood, Elementary and Secondary Education	1
Prepared statement of	3
Submissions for the record:	
Prepared statement of Eva L. Baker, distinguished professor, director, UCLA, Center for Research on Evaluation, Standards, and Student Testing (CRESST)	6
Prepared statement of Linda Darling-Hammond, Charles E. Ducommun professor, Stanford University School of Education	12
Statement of Witnesses:	
Finn, Chester E., Jr., president, Thomas B. Fordham Institute and chairman, Koret Task Force on K–12 Education, Hoover Institution, Stanford University	43
Prepared statement of	45
Jennings, Jack, president, Center on Education Policy	24
Prepared statement of	26
Johnson, Dr. Carol, superintendent of the Memphis City Schools	37
Prepared statement of	39
Melmer, Hon. Rick, South Dakota Secretary of Education	29
Prepared statement of	31
Straus, Hon. Kathleen N., president, Michigan State Board of Education	33
Prepared statement of	35

REAUTHORIZATION OF THE ELEMENTARY AND SECONDARY EDUCATION ACT: CURRENT AND PROSPECTIVE FLEXIBILITY UNDER NO CHILD LEFT BEHIND

Thursday, June 7, 2007
U.S. House of Representatives
Subcommittee on Early Childhood,
Elementary and Secondary Education
Committee on Education and Labor
Washington, DC

The subcommittee met, pursuant to call, at 2:30 p.m., in Room 2175, Rayburn House Office Building, Hon. Dale Kildee [chairman of the subcommittee] Presiding.

Present: Representatives Kildee, Kucinich, Davis of California, Payne, Sarbanes, Loebsack, Hirono, Hare, Woolsey, Castle, Hoekstra, Ehlers, Biggert, Bishop of Utah, Wilson, Kuhl, and McKeon.

Staff Present: Tylease Alli, Hearing Clerk; Alice Cain, Senior Education Policy Advisor (K-12); Adrienne Dunbar, Legislative Fellow, Education; Lloyd Horwich, Policy Advisor for Subcommittee on Early Childhood, Elementary and Secondary Education; Lamont Ivey, Staff Assistant, Education; Thomas Kiley, Communications Director; Danielle Lee, Press Outreach Assistant; Jill Morningstar, Education Policy Advisor; Joe Novotny, Chief Clerk; Alex Nock, Deputy Staff Director; Lisette Partelow, Staff Assistant, Education; James Bergeron, Minority Deputy Director of Education and Human Services Policy; Robert Borden, Minority General Counsel; Kathryn Bruns, Minority Legislative Assistant; Steve Forde, Minority Communications Director; Taylor Hansen, Minority Legislative Assistant; Victor Klatt, Minority Staff Director; Chad Miller, Minority Professional Staff; Susan Ross, Minority Director of Education and Human Services Policy; Linda Stevens, Minority Chief Clerk/Assistant to the General Counsel; and Cameron Coursen, Minority Assistant Communications Director.

Chairman KILDEE. A quorum being present, the hearing of the subcommittee will come to order. First of all, I love being in the majority, but on a day like this the minority seems rather attractive. I have got three things going on and you are one of the major ones here, so I literally ran over from the other building. Good to see all of you. You are all friends of mine and I have known most of you for many, many years and welcome you.

Pursuant to Committee Rule 12(a) any member may submit an opening statement in writing which will be made part of the permanent record. I would now recognize myself followed by Ranking Member Castle for opening statements.

I am pleased to welcome my fellow subcommittee members, the public and our witnesses to this hearing on

Reauthorization of the Elementary and Secondary Education Act. We very often use the term "expert witnesses." today the table speaks for itself. We have expert witnesses here today.

I would like to talk about the current and prospective flexibility of No Child Left Behind. This is the eighth hearing that this subcommittee has held on No Child Left Behind this year. The full committee of course has held many hearings as well. I think it is safe to say there is no subject we hear more about than flexibility.

As written, the law provides for certain flexibilities and the Department of Education has provided others. But many State and local educators have told us that while they strongly support the law's goals and the discussion about accountability that it has fostered, better flexibility would help them to reach those goals. I take their comments very seriously because I always have believed that education is a local function, and Jack Jennings used to hear me say this all the time years ago, a State responsibility but a very important Federal concern.

As our society and the world have become more mobile and more interconnected that national concern has grown. People educated in one State wind up in another and we are competing in a global economy and education and training will give us the edge in that competition. And regardless of where in the United States students live, they ultimately will compete with students from around the world. But a greater national concern does not mean less emphasis on State responsibilities and local functions.

And so I look forward to hearing from our witnesses and how flexibility under No Child Left Behind has been implemented and how we can improve that flexibility. Their testimony as well as the countless conversations that I know each member has had with educators and parents in their district and here in Washington will play a critical role in the committee's efforts to understand how we can best help to provide every student with a world class education, a goal we all share.

I also look forward to hearing from Mr. Jennings about his center's recent studies on trends and student achievement since No Child Left Behind took effect. The title of that study, "Answering the Question That Matters Most, Has Student Achievement Increased Since No Child Left Behind," is well chosen since in the end the point of all this is student achievement.

Of course one factor that has not increased enough under No Child Left Behind has been funding. We owe it to our children to ensure that their schools have the resources and support to provide them with the education they need and deserve. Since 2002, Congress and the President have underfunded No Child Left Behind by $56 billion. And the President's proposed budget for 2008 would underfund the law by another $15 billion for a total of $71 billion. However, I am hopeful that with the changes in Washington this year we will start to do better. Our budget resolution calls for that.

But I look forward to continuing to work with my ranking member, Mr. Castle, our full committee chairman and ranking member, Mr. Miller and Mr. McKeon, and with all the members of the committee on a bipartisan reauthorization of No Child Left Behind this year. And I yield to my friend and the ranking member, Governor Castle.

[The prepared statement of Mr. Kildee follows:]

Prepared Statement of Hon. Dale E. Kildee, Chairman, Subcommittee on Early Childhood, Elementary and Secondary Education

I'm pleased to welcome my fellow subcommittee members, the public, and our witnesses, to this hearing on "Reauthorization of the Elementary and Secondary Education Act: Current and Prospective Flexibility Under No Child Left Behind."

This is the eighth hearing that this subcommittee has held on No Child Left Behind this year.

The full committee, of course, has held many hearings as well.

And I think it's safe to say that there is no subject we hear more about than flexibility.

As written, the law provides for certain flexibilities, and the Department of Education has provided others.

But, many state and local educators have told us that while they strongly support the law's goals and the discussion about accountability that it has fostered, better flexibility would help them to reach those goals.

I take their comments very seriously, because I always have believed that education is a local function, a state responsibility and a national concern.

As our society and our world have become more mobile and more interconnected, that national concern has grown.

Many students from the home state of each member of this subcommittee will one day move to the home states of every other member of the subcommittee.

And regardless of where in the united states students live, they ultimately will compete with students from around the world in the global economy.

But, a greater national concern does not mean a lesser emphasis on state responsibilities and local functions.

And so, I look forward to hearing from our witnesses on how flexibility under NCLB has been implemented and on how we can improve that flexibility.

Their testimony, as well as the countless conversations that I know each member has had with educators and parents in their district and here in Washington, will play a critical role in the committee's efforts to understand how we can best help to provide every student with a world-class education—a goal we all share.

I also look forward to hearing from Mr. Jennings about his center's recent study on trends in student achievement since NCLB took effect.

The title of that study "Answering the Question That Matters Most: Has Student Achievement Increased Since No Child Left Behind?" is well chosen, since in the end, the point of all this is student achievement.

Of course, one factor that has not increased enough under No Child Left Behind has been funding.

We owe it to our children to ensure that their schools have the resources and support to provide them with the education they need and deserve.

Since 2002, Congress and the President have underfunded No Child Left Behind by $56 billion.

The president's proposed budget for 2008 would underfund the law by another $15 billion, for a total of $71 billion.

However, I am hopeful that with the changes in washington this year, we will start to do better.

But, I look forward to continuing to work together with my ranking member, Mr. Castle, our full committee chairman and ranking member, Mr. Miller and Mr. McKeon, and with all the members of the committee, on a bipartisan reauthorizaton of NCLB this year.

I now yield to Ranking Member Castle for his opening statement.

Mr. CASTLE. Thank you, Mr. Chairman. I would like to thank you for holding today's hearing. As far as your comments and et cetera being in the majority and the minority, having a busy extremely day myself in the minority, I don't think it makes much

difference in terms of how busy we are around here. And if you want to switch at any time just let me know and we'll try to work out political party differences.

This is the latest in a series of hearings on No Child Left Behind. Some of you out in this audience have followed that. One thing nobody can ever accuse this committee of is not having sufficient hearings on No Child Left Behind. We seem to have had them almost on a weekly basis last year and this year as we get ready to go up to the reauthorization.

We obviously welcome all the witnesses here today. I look forward to your testimony. No Child Left Behind represents a comprehensive overhaul of the Elementary and Secondary Education Act and was passed to adjust the achievement gap that exists between low income and minority students and their more affluent peers.

Earlier this week the Center on Education Policy released the results of a study of trends and State test scores since No Child Left Behind became law, and I am pleased to have the President of CEP with us today to discuss these findings in greater depth because the results are promising. For instance, students are doing better on State reading and math tests since the No Child Left Behind was enacted. And the greatest is being made on elementary school math tests. While these findings are encouraging additional flexibility for States in school districts which have differing priorities will help a greater number of their students reach proficiency.

Flexibility is already a key element within No Child Left Behind. Currently individual States are given the flexibility to determine a variety of student achievement factors, including the definition of proficiency, the starting point for progress measurement and the amount of progress that must be made from year to year. States also have the flexibility to develop their own tests to determine if existing teachers should be deemed highly qualified.

Additionally, No Child Left Behind includes a number of new and existing flexibility provisions that aim to increase the ability of States and local school districts to use Federal assistance to meet their own priorities, including the transferability of Title I funds, school-wide programs to improve services to all students and State and local flexibility demonstration programs.

In addition to the valuable information each of our witnesses will share with us today about flexibility provisions under No Child Left Behind, Mr. McKeon, the senior Republican of the full committee recently introduced the State and Local Flexibility Improvement Act. This legislation builds upon the success of No Child Left Behind by further strengthening flexibility under the laws for States and local school districts. As an original cosponsor of this legislation, I believe this legislation gives States and local school districts the freedom to target Federal resources to best serve the needs of their students, while maintaining strong accountability standards to measure would allow States to waive certain statutory or regulatory requirements under law, consolidate Federal education programs and use an alternative method for making allocations in local school districts instead of their current formula if their new proposal targets funds more effectively to those areas with high concentrations of low income families, measure indi-

vidual student growth, including through well-designed growth models, expand the poverty threshold for school-wide programs and most notably allow States and school districts to transfer 100 percent of their Federal programs within certain programs, up from 50 percent under current law, into the Title I program.

As I said, I believe strongly in No Child Left Behind. The importance of closing the achievement gap cannot be overstated and I believe Mr. McKeon's bill will help States and local school districts close that gap even more quickly. I look forward to hearing from today's witnesses about flexibility under No Child Left Behind as Congress begins to reauthorize this law, and I thank you all very much.

Thank you, Mr. Chairman. I yield back.

[The prepared statement of Mr. Castle follows:]

Prepared Statement of Hon. Michael N. Castle, Ranking Minority Member, Subcommittee on Early Childhood, Elementary, and Secondary Education

Good Afternoon. I'd like to thank Chairman Kildee for holding today's hearing, the latest in our series of hearings on the No Child Left Behind Act. I also would like to welcome our witnesses and thank you all for being here to testify today.

NCLB represents a comprehensive overhaul of the Elementary and Secondary Education Act and was passed to address the achievement gap that exists between low-income and minority students and their more affluent peers.

Earlier this week, the Center on Education Policy released the results of a study of trends in state test scores since NCLB became law. I am pleased we have the president of CEP with us today to discuss these findings in greater depth because the results are promising. For instance:

- Students are doing better on state reading and math tests since the No Child Left Behind Act was enacted
- And the greatest progress is being made on elementary school math tests.

While these findings are encouraging, additional flexibility for states and school districts, which have differing priorities, will help a greater number of their students reach proficiency.

Flexibility is already a key element within NCLB. Currently, individual states are given the flexibility to determine a variety of student achievement factors, including the definition of proficiency, the starting point for progress measurement, and the amount of progress that must be made from year to year. States also have the flexibility to develop their own tests to determine if existing teachers should be deemed highly qualified.

Additionally, NCLB includes a number of new and existing flexibility provisions that aim to increase the ability of states and local school districts to use federal assistance to meet their own priorities including the transferability of Title I funds, school-wide programs to improve services to all students, and state and local flexibility demonstration programs.

In addition to the valuable information each of our witnesses will share with us today about flexibility provisions under No Child Left Behind, Mr. McKeon, the Senior Republican of the Full Committee, recently introduced the State and Local Flexibility Improvement Act. This legislation builds upon the success of NCLB by further strengthening flexibility under the law for states and local school districts.

As an original cosponsor of this legislation, I believe this legislation gives states and local school districts the freedom to target federal resources to best serve the needs of their students.

While maintaining strong accountability standards, the measure would:

- Allow states to waive certain statutory or regulatory requirements under law, consolidate federal education programs, and use an alternative method for making allocations to local school districts instead of their current formula if their new proposal targets funds more effectively to those areas with high concentrations of low-income families;
- Measure individual student growth, including through well-designed growth models;
- Expand the poverty threshold for school-wide programs; and
- Most notably, allow states and school districts to transfer 100 percent of their federal funds within certain programs, up from 50 percent under current law, into the Title I program

As I have said, I believe strongly in No Child Left Behind. The importance of closing the achievement gap cannot be overstated, and I believe Mr. McKeon's bill will help states and local school districts close that gap even more quickly. I look forward to hearing from today's witnesses about flexibility under NCLB as Congress begins to reauthorize this law.

Thank you.

Chairman KILDEE. Thank you, Governor. Without objection, all members will have 7 calendar days to submit additional materials or questions for the hearing record.

[The prepared statement of Eva L. Baker, submitted by Mr. Kildee, follows:]

Prepared Statement of Eva L. Baker, Distinguished Professor, Director, UCLA, Center for Research on Evaluation, Standards, and Student Testing (CRESST)

I thank Chairman Miller for the opportunity to provide testimony.

Below is a summary of the main ideas of testimony related to No Child Left Behind*

1. Reduce the number of subject matter tests required, keeping the focus on elementary education.
2. If multiple measures are to be used, involving student performance, consider they are compensatory rather than conjunctive use in elementary school, so they do not add risk to schools AYP.
3. Provide incentives for States to develop and use valid tests that assess students' ability to transfer and apply knowledge. Encourage flexibility to determine the relative value that should be accorded broadly sampled knowledge compared to evidence of depth of study.
4. Consider multiple studies of the validity of measures now in use, sampling over States.
5. Provide incentives for States to experiment and validate "opportunity to learn" in classroom practice, to determine the breadth and range of teaching.
6. Conduct validity studies related to the performance and types of instruction experienced by groups such as English learners and low scorers on State Examinations.
7. Change achievement criteria for secondary school by immediately authorizing the use of qualifications for students to be completed by their graduation. These qualifications are not single tests, but provide integrated choices for students to acquire and demonstrate expertise in subject matter, in tasks related to societal needs, e.g., environmental studies, or in topics such as the performing arts. Qualifications may be relevant to future work or postsecondary study but not restricted to them. They must count in accountability formulae. (Note that excellent examples may be seen in Web sites of education ministries in other countries, for instance, in the New Zealand system: See *http://www.nzqa.govt.nz/qualifications/index.html)*
8. Revise the Adequate Yearly Progress requirements to permit more flexibility. Flexibility may be attained by recording performance on subsets of examinations, that could be taken at various times during the year, by considering periods longer than a year or two for the computation of sanction-related AYP performance, and by reducing the artifacts (such as number of groups) that influence AYP.

Take a moment and consider the following question. What needs to change in your life? I have a confession to make. You may be surprised to learn that I want a life with more balance. Over the years, I've asked many AERA members to name a professional who lives a balanced life and always with the same result. Long stares and no names. For most of us, however, balance is victim of competing values, looming obligations, unfulfilled ambitions, and in my case, wearable media. I think it no accident that balance is a metaphor for justice. The lack of balance takes different forms but is found across the entire economic spectrum—for all groups and all kinds of people—for the old, for the young, and relevant here, for those in schools.

---

*Note: In the interest of brevity, I will not discuss aspects of the legislation that may be regarded as positive and rather focus on areas that I believe need attention. I attach a draft of my AERA presidential address, which spells out a justification for my recommendations. Because the presentation is intended to be oral, it does not contain the research citations you may want. If you desire citations, I can provide them upon my return to Los Angeles the week of April 15.

In particular, U.S. adults consistently report in surveys that their lives are out of whack. I think we have inadvertently projected this imbalance onto the lives of schools.

So this talk—The End(s) of Testing—is about balance—what we have now, and ways to move to a new equilibrium—where we reconnect achievement to learning, equity to more than equal test scores, and students to their own paths. With advance apologies, I'll use a few examples from other countries and liberally incorporate ideas from research here and around the world, without oral citation because of time.

Consider Jerome Bruner's statement "Life in culture is, then, an interplay between the versions of the world that people form under its institutional sway and the versions of it that are products of their individual histories." Accountability tests have swung education strongly toward institutional goals and away from those of the individual. We need common goals and measures to identify and fix inadequacies and inequities. But these accountability indicators have dramatic effects on teaching and learning. Cramped by requirements, we harden instruction, drop electives, and shorten time for in-depth engagement. Students' voices and choices are fainter. Since the world is flat in testing, too, test results everywhere strongly affect classroom realities and public perceptions of learning and schooling.

Balance is a goal, and a fundamental design principle essential to quality itself. Balance invariably involves trade-offs, so what do we give up to gain educational balance? Let's start with what we have.

On Current Standards and Accountability Measures

Daily, or more often, researchers and education writers bemoan the technical shortfalls of tests now used in accountability, contending that the content of these tests is off, their sampling is odd, and deep cognition is lost. Nor do the institutional tests seem to matter much to students. I add to that litany that these tests virtually ignore learning research and transfer of knowledge, that is, the application of learning to something other than to another test. Tests only dimly reflect in their design the results of research on learning, either of skills, subject matter, or problem solving. These test properties matter to researchers, but rarely to others. Many professionals, and I would guess, most of the public, don't make fine distinctions. They see tests on the same topic as interchangeable, and a high score on any test as sufficient evidence of learning.

To my mind, the evidential disconnect between test design and learning research is no small thing. Think about it. It means, at worst, the tests may not actually be measuring the learning for which schools are responsible, thus gutting the basic tenet of the accountability compact. So set aside learning-based design, and ask, "How well do any of our external tests work?" The answer is that we often don't know enough to know. We have little evidence that tests are in synch with their stated or de facto purposes, or that their results lead to appropriate decisions. In other words, we act as if tests were valid, in the face of weak or limited evidence. Notwithstanding, we make heavy and far-reaching decisions about schools and students, talk about gaps, and applaud progress. This excitement takes place with only fragments of evidence called for by the Test Standards of AERA, NCME, and APA. Educational researchers conduct validity studies, when possible, and have called for high-quality, evidence-based tests for years. We know the importance of right inferences by teachers, administrators, and the government. We should be at least queasy about the quality of the interpretations, the meaning of gaps and improvement, and the resultant classifications of schools. Yet, test validity languishes as a largely unexamined, prior question because of inexorable schedules and budget constraints. With tests of uncertain validity, adequate yearly progress (AYP), value-added, or other growth modeling analyses will have limited meaning in accountability interpretations.

So how did we get here? The wave of U.S. reform was stimulated, in part, by lackluster performance on international comparisons more than two decades ago. Not surprisingly the reform plan was to follow the international lead and design a quasi-national system of standards and assessments. Despite awareness of huge differences in context and traditions (our 50-State autonomy in education, distributed curricula, independent teacher education institutions, and waning respect for those working in education), sets of State and Federal legislation enabled state standards and related tests.

We see once more how tests exert power. Research in both schools and laboratories shows that testing improves learning of what is tested, even in the absence of feedback, and that with feedback, learning improves more. When results are linked to sanctions, teaching moves to conform to the content boundaries of the tests. Because topics on tests are represented inequitably, to avoid sanctions it is

efficient to practice TEST-LIKE items. Practice them a lot. Dan Koretz calls this score inflation. You may call it smart. In either case, it is likely that test-specific content and formats are learned through test practice at the expense of the intended content and skill domains promised by standards. But some learning takes place. Now a new kind of test anxiety uses up all the instructional oxygen in the classroom.

Coping Strategies and "Accountabalism"

While accepting the importance of accountability, educators in the U.S. and internationally have anticipated negative consequences of hyper test consciousness. David Weinberger, in the Harvard Business Review, describes a process called "Accountabalism" (think of Hannibal Lector with a checklist). Weinberger illustrates the consequences of a repair mentality that deals serially with each successive difficulty in accountability systems as if we could perfect the system. Although not focused on schools, his analysis of accountability suggests that a chronological, piecemeal fix does really not improve systems. Over time, repeated patching and spackling consumes all remaining flexibility. Sound familiar?

Mitigations

In attempting to head off some of the negative side effects of accountability while still accepting its value, educators advocate six tactics. The first tactic is to add more or different measures or indictors to accountability, in shorthand, to use "multiple measures" (although the original intent of the term was to allow students different ways to show their competence). One approach adds a periodic test, say, of history, with rotating content to the system, to be incorporated into the accountability calculation.

A second type calls for measures of opportunity to learn or OTL as part of accountability models. OTL was early thought of as a check both on the fair access to test-relevant instruction and as a shield to protect non-tested content and activities, such as different courses. Scalable, valid methods for routinely assessing inside-the-classroom OTL remain an ongoing research quest. When there are efficient, replicable, and valid procedures for its measurement, the role of OTL is clear in accountability.

"Having tests worth teaching to," despite its grammatical failure, is the mantra of the third approach, performance assessment. Change the test and legitimate test preparation. Serve good rather than evil. Performance assessment calls for multistepped activities, usually imbued with some realism. More significantly, performance assessment design originally reflected research from teaching, from cognitive psychology, and from subject matter learning. In its heyday, models from abroad were emulated, and tasks were sometimes directly copied. After handsome financial support, popularity waned and only vestiges of performance assessment, such as written composition, can be found on mandated tests—and in higher education. Why the demise? The list includes lack of political acceptance and structural integration, feasibility, technical quality, struggles in longitudinal use, credibility (because of overselling), and cost.

These three examples of multiple measures, unless done wisely, offer schools the scary option of more ways to fail.

A fourth and currently "hot" mitigating tactic is formative assessment. Assessment to support learning has been continuously in play since writing by Vygotsky, Skinner, and Lindquist, and before. Black and Wiliam, James, and Heritage, all from abroad, by the way, now invoke a process where teachers ask penetrating questions, pinpoint errors, give insightful and timely feedback, and use innovative teaching to help students. Not surprisingly, has been the entry of commercially supplied benchmark or interim tests are in use, carrying the rationale for formative assessment, if not its spirit. For use at intervals during instruction, interim or benchmark tests give previews of coming results, based on test segments similar to the main assessment. The validity of these measures is rarely cited, other than similar items predict results on similar items. Although, their results are intended to influence teaching, because of strict instructional pacing, some teachers have little or no time to use these interim results and must move on. Even with adequate time, teachers need expertise in their subject matter understanding of their students, and fluency in alternate ways of teaching. Teachers with these skills are in short supply where they are needed most.

A fifth tactic limits the number of standards to be tested by setting clear priorities. A less-is-more stance. The Commission on Instructionally-Supportive Testing, prepared a report arguing that fewer, but more powerful standards will lead to valid measurement, more coherent teaching, and deeper student learning. Their proposal directly confronts the political bargains that underlie expansive lists of standards.

A sixth approach is technology, and it is my favorite. I'm the one, after all, with a Blackberry and a Treo, waiting anxiously for the iPhone, to complement my 4 iPods, 5 computers, and the Slingshot TiVo. Intelligent tutoring systems show us how to link assessment and learning models. Yet, technology-linked tests have not yet stepped up to early expectations. In large-scale testing, they serve efficiency without capitalizing on their potential to leverage higher fidelity experience (like simulations), or exploit students' ease with technology.

As a way forward to revitalize performance assessment, technology offers some hope to find more effective ways to assess validity with regard to learning, and to reduce the cost of design and use. There is excellent progress in computer scoring of open-ended written responses; with optical character reading, students' handwritten work can be scanned and directly marked by computer. Speech recognition technologies process oral language, so discourse of learners of all ages and language backgrounds can be partly analyzed. Expect to see more computer games and virtual worlds as assessment contexts, as well as assessment embedded in common devices such as cell phones and game platforms—first to reduce gaps in computer access, and second to leverage motivation with fun and the familiar. A second use of common devices is to permit "anytime, any where" or "just-in-time" assessment. To speed and improve assessment design, look to computer-assisted assessment authoring systems that can improve assessment design with built-in knowledge to assist users to make, mark, and manage tests.

International Comparisons

If our showing on international comparisons started us down this path, what do recent international tests tell us? The Organisation for Economic Co-Operation and Development (OECD) has fielded the Programme for International Student Assessment (PISA), examinations that include cross-curricula content and the application of cognitive skills, such as problem solving. Their results are worth reflection. The vertical bars show great performance variation between the best and poorest performing students in most countries on some 21st century skills. This variation applies almost independently of average score, except for size and homogeneity of national populations. Why? Are current curricula not sufficiently attuned to future requirements, are students not taught to transfer learning, or to apply cognitive skills to new situations? Don't all students deserve to learn these skills?

CRESST POWERSOURCE©—A U.S.-International Collaboration

I will now briefly touch on our own integration of a balanced formative assessment system, with design based on learning research, including schema development, explanation, narrative, and transfer. Based on about 40 earlier studies, POWERSOURCE©, funded by our Institute of Education Sciences (IES) Center award, is our major experimental project in middle-school, pre-algebra, and rests on multiple interim assessments of problem solving and explanation. POWERSOURCE© development begins with analyzing cognitive demands, then representing the relationship of these intellectual skills in a content ontology. The content includes "big Mathematics ideas," the power principles, which we construe as schema, to be applied flexibly across topic and problem type. The ontology controls sampling and sequence, and also serves as a learning performance aid to help teachers and students map where they are during instruction. Our assessments are based on research by Sweller, Mayer and their colleagues on schema acquisition and worked examples, on explanation studies (by Chi and by our own team), and on transfer research (by Bjork and Holyoak and others). Assessments are embedded in a kid-friendly narrative theme. We predict in this 3-year longitudinal study improved teacher content and pedagogical knowledge. We expect superior student performance on state assessments, and, because of their schema acquisition, higher performance on transfer tasks drawn both from the Key Stage 3 maths tests in England and from the PISA examination. This experiment is being replicated in Korea at that country's own expense, to generate comparative information about the use of learning research in designing assessment. With POWERSOURCE©, we illustrate assessment balance with a method that transfers to different subjects and ages of students.

The International Stage of Examinations

An obvious question given their initial impetus, is how closely do U.S. tests resemble examinations from abroad, such as those in New Zealand, or Hong Kong, or Finland? The quick answer is they don't, despite, in some countries, the use of similar test formats. But response format aside, most national testing differs substantially from the U.S. versions. These countries make far greater investment in integrating curriculum, professional development, and assessment, thus removing the need for post hoc rationalization of alignment. Extensive care is given to the

content of the test questions, and to model answers used to guide marking. (Of course, national testing systems are themselves subject to continuing and sometimes scathing debate.)

Visit Web sites of education departments and ministries around the world and you will see great variety. Notice the range existing within many examination systems. Whether workforce or university paths, secondary school students in other countries have many options. They select from a wide range of course-based examinations, in some countries close to 50 choices. Some options look like the familiar Advanced Placement or International Baccalaureate courses; other choices encourage high school students to pursue interests unconnected to traditional university or work requirements, for instance, in the visual and performing arts or environmental studies.

Before the Future

Before I sketch a way to attain balance for students and schools, let's return to current accountability practice. First, the tests—I say let's leave them alone. They are resilient, embedded in our traditions, and changes to them are always temporary (they snap back like a rubber band), with at best trivial residue. To attain better balance, I'd have fewer tests, locating them during skill building in elementary schools. Even if their numbers were reduced, we still must demand studies of validity inferences and the tests' multiple purposes. With current accountability systems, in the U.S., I say fix the elements of AYP to minimize harm, by considering the raw probabilities of failure related to numbers of groups by partitioning of tests into passable components, and by aggregating results in larger units than a year or two. It is appealing to consider new indicators or multiple measures to balance testing, but their integration in accountability numbers must avoid artifacts that put schools inappropriately at risk.

Change

"The scientific revolution that began 300 years ago has accelerated exponentially. It is moving so fast that the spread of knowledge defines our times."

"There will be more change in the next ten years than in the previous 100 years."

Fareed Zakaria, 2006, 2007.

"640K ought to be enough for anybody."

Bill Gates, 1981.

If our accountability system stays locked on producing only common (and needed) achievement, we will continue to open wider the chasm between what is sanctioned and what student futures demand. So the matter of balance returns. Can we have balance and deal with the students we have now, without wishing they were different? We can't wait until all high school students have developed proficient language skills, attended innovative preschools, or learned number facts from amazing third-grade teachers. So here and now, can we offer students options better suited to a new and changing work environment, to a life in a real-time society with amped-up connectivity? And can we teach them in a way that models flexibility and problem solving? Researchers Glaser, Simon, O'Neil, and Perez, among others, have listed skill-sets for a future brimming with choices and expectations:

- Adaptive problem-solving abilities;
- Capacity to assess and respond to risk;
- Managing distraction, and giving mindful, rotating attention to tasks;
- Solitary work, with self management;
- Changeable roles in real or virtual teams and groups.

Posit these as a beginning of new 21st century definition of educational quality. Start in secondary school where we all have unresolved and growing problems. There we should create rapidly a system of Qualifications to reflect 21st century needs, to be available to all students, whatever their status on standards-based tests. My image of a Qualification is validated accomplishment, obtained inside or outside of school. A Qualification means just that: at different levels of challenge, that a student has attained a certified, trusted accomplishment. Warmed-over performance assessment? I think not. Each Qualification is not a new test, but an integrated experience with performance requirements. It might look like a course, or a collection, or a musical or sports performance. Some Qualifications may demand shorter, intense involvement like securing a certification in CPR, or network management. They come aligned, with integrated goals, tasks, learning experiences, criteria and tests. Merit badge metaphors, like one proposed many years ago by Al Shanker, may help you get the idea. It is a truth in advertising approach; instead of a set of "scores," a student possesses demonstrated accomplishments.

Qualifications seem to have many benefits. They give clear venues for high school students to improve skills, apply and adapt knowledge, and acquire new learning. They support and credit emerging maturity to develop desired personal expertise.

To presage construction, an ontology of Qualifications will give structure to their features. Let me illustrate an ontological dimension of Qualifications related to learning. For example, as design criteria, each Qualification could represent:

a. Complex problem solving and reasoning;
b. Flexibility and adaptive performance;
c. Rich knowledge base;
d. Schema or principle learning;
e. Metacognition and self-monitoring;
f. Communication, either explanatory or interactive.

The Qualifications chosen by students should span a range of personal goals, and with any luck help them to develop a passion in at least one area.

What could Qualification topics include? Certainly disciplinary inquiry, exemplified in different ways, perhaps in science fairs, competitions, or projects; then there are the performing or visual arts, community service, healthcare, interning in business, environmental studies, teaching the young, mentoring peers, and helping the old.

Who will offer these options? Some may be available in high schools, in after-school programs, in community colleges, or in universities. I expect that the Web delivery, with its entrepreneurial brainpower and finance will be a powerful and welcome source of creative Qualifications. Public and private spaces, such as museums, businesses, government, advocacy, private and service organizations, can help too.

How to begin? Our international colleagues have wonderful Qualifications that we can adapt. If we want to take this reform path, it will be a national challenge of commitment. At the very least, the expectations and rhythms of secondary school will change. Although Qualifications will need to be understood, energized and partly managed by teachers, expertise outside of school will matter, too.

Is this rebalancing of secondary school curricula feasible? Yes, if four benchmarks are met within a reasonable time.

First, the framework(s) for these Qualifications must reside in one or more politically credible national or State organizations. Achieve, Inc. has a great start in its American Diploma Project with more than half of the states participating. Or see the State Department of Instruction in North Carolina for another example. Benchmark 2 is quality control. What evidence of quality is needed for a Qualifications adoption? Forms of evidence could include empirical comparisons or expert judgment.

Benchmark 3, acceptance by business and universities in their selection processes, is essential for the Qualifications to matter to students. Benchmark 4, and of extraordinary importance, is that Qualifications be made a part of State and Federal accountability systems and linked to other politically powerful structures. With that role, we need protections against phony performance, plagiarism, and overzealous parental help.

How fast? As a lover of 10-year funded research grants, I'm sorry to say our timelines in this area must be unusually brief. No years of trials and pilots, but concurrent validity studies. By starting with examples and data from other countries or adopting existing Qualifications from our own business and arts sectors, we can have a pool of 40 Qualifications available by 2008, but only as a start. We should use systematic R&D models from other national ventures and invest in parallel development of Qualifications in different areas. The federal government, along with the private sector, can be a financial catalyst. A modified NCLB should phase in Qualifications NOW.

Unless we begin right now to fix educational balance, with a strong focus on secondary schools, we will have stark questions to answer. Who decided to triage large numbers, and particular groups of secondary school students? Why don't we care that many will not find a life of contribution and meaning? A goal of balance in the revitalization of secondary schooling will develop generations of students to be far better prepared than we are.

Research Agenda

The research agenda is long, and at this hour our attention is short (I'll make a complete list in the written version of the paper). Yet, opportunities traverse theories of learning and validity, explorations of equity, individual and group differences, definitions of adaptability, studies of transfer within and across fields, expertise, narrative, efficacy, self-management, and motivation. For both students and

teachers. We can look at second or third chances, mobilizing community and business interests, and of course, cost.

So to return to the beginning, can we find ways to get balance into the schools and promote a different quality of learning—within the essential framework of accountability? Can we take on the high school substantively? Let's cross over to a new path, built on previous research, and reinvest in learning, where accomplishments come with validity, and the balance is redressed between what we think high school students need and what they think they need. The paths of Qualifications shift attention from school work to usable and compelling skills, from school life to real life. With pride our students can assemble their unique collection of Qualifications, to show to their families, adults in university and workforce, and to themselves. With collaboration of both the international community and our own communities, we can enable education to prepare our students far better for the future. As you know, when we are balanced, each of us is able to move easily in a range of directions. So with balance, and help from the world community, our students succeed, and fulfill their not yet imagined dreams.

[The prepared statement of Linda Darling-Hammond, submitted by Mr. Kildee, follows:]

Prepared Statement of Linda Darling-Hammond, Charles E. Ducommun Professor, Stanford University School of Education

I thank Chairman Miller and the members of the Committee for the opportunity to offer testimony on the re-authorization of ESEA, in particular the ways in which we measure and encourage school progress and improvement. My perspective on these issues is informed by my research, my work with states and national organizations on standards development, and my work with local schools. I have studied the implementation of No Child Left Behind,[1] as well as testing and accountability systems within the United States and abroad.[2] I have also served as past Chair of the New York State Council on Curriculum and Assessment and of the Chief State School Officers' INTASC Standards Development Committee. I work closely with a number of school districts and local schools on education improvement efforts, including several new urban high schools that I have helped to launch. Thus, I have encountered the issues of school improvement from both a system-wide and local school vantage point.

I am hopeful that this re-authorization can build on the strengths and opportunities offered by No Child Left Behind, while addressing needs that have emerged during the first years of the law's implementation. Among the strengths of the law is its focus on improving the academic achievement of all students, which triggers attention to school performance and to the needs of students who have been underserved, and its insistence that all students are entitled to qualified teachers, which has stimulated recruitment efforts in states where many disadvantaged students previously lacked this key resource for learning.

The law has succeeded in getting states, districts, and local schools to pay attention to achievement. The next important step is to ensure that the range of things schools and states pay attention to actually helps them improve both the quality of education they offer to every student and the quality of the overall schooling enterprise. In order to accomplish this, I would ask you to actively encourage states to:

- Develop accountability systems that use multiple measures of learning and other important aspects of school performance in evaluating school progress;
- Differentiate school improvement strategies for schools based on a comprehensive analysis of their instructional quality and conditions for learning.

Why Use Multiple Measures?

There are at least three reasons to gauge student and school progress based on multiple measures of learning and school performance:

- To direct schools' attention and effort to the range of measures that are associated with high-quality education and improvement;
- To avoid dysfunctional consequences that can encourage schools, districts, or states to emphasize one important outcome at the expense of another; for example, focusing on a narrow set of skills at the expense of others that are equally critical, or boosting test scores by excluding students from school; and
- To capture an adequate and accurate picture of student learning and attainment that both measures and promotes the kinds of outcomes we need from schools.

Directing Attention to Measures Associated with School Quality

One of the central concepts of NCLB's approach is that schools and systems will organize their efforts around the measures for which they are held accountable. Because attending to any one measure can be both partial and problematic, the concept of multiple measures is routinely used by policymakers to make critical decisions about such matters as employment and economic forecasting (for example, the Dow Jones Index or the GNP) and admission to college, where grades, essays, activities, and accomplishments are considered along with test scores.

Successful businesses use a "dashboard" set of indicators to evaluate their health and progress, aware that no single indicator is sufficient to understand or guide their operations. This approach is designed to focus attention on those aspects of the business that describe elements of the business's current health and future prospects, and to provide information that employees can act on in areas that make a difference for improvement. So, for example, a balanced scorecard is likely to include among its financial indicators not only a statement of profits, but also cash flow, dividends, costs and accounts receivable, assets, inventory, and so on. Business leaders understand that efforts to maximize profits alone could lead to behaviors that undermine the long-term health of the enterprise.

Similarly, a single measure approach in education creates some unintended negative consequences and fails to focus schools on doing those things that can improve their long-term health and the education of their students. Although No Child Left Behind calls for multiple measures of student performance, the implementation of the law has not promoted the use of such measures for evaluating school progress. As I describe in the next section, the focus on single, often narrow, test scores in many states has created unintended negative consequences for the nature of teaching and learning, for access to education for the most vulnerable students, and for the appropriate identification of schools that are in need of improvement.

A multiple measures approach that incorporates the right "dashboard" of indicators would support a shift toward "holding states and localities accountable for making the systemic changes that improve student achievement" as has been urged by the Forum on Education and Accountability. This group of 116 education and civil rights organizations—which include the National Urban League, NAACP, League of United Latin American Citizens, Aspira, Children's Defense Fund, National Alliance of Black School Educators, and Council for Exceptional Children, as well as the National School Boards Association, National Education Association, and American Association of School Administrators—has offered a set of proposals for NCLB that would focus schools, districts, and states on developing better teaching, a stronger curriculum, and supports for school improvement.

Avoiding Dysfunctional Consequences

Another reason to use a multiple measures approach is to avoid the negative consequences that occur when one measure is used to drive organizational behavior.

The current accountability provisions of the Act, which are focused almost exclusively on school average scores on annual tests, actually create large incentives for schools to keep students out and to hold back or push out students who are not doing well. A number of studies have found that systems that reward or sanction schools based on average student scores create incentives for pushing low-scorers into special education so that their scores won't count in school reports,[3] retaining students in grade so that their grade-level scores will look better,[4] excluding low-scoring students from admissions,[5] and encouraging such students to leave schools or drop out.[6]

Studies in New York,[7] Texas,[8] and Massachusetts,[9] among others, have showed how schools have raised their test scores while "losing" large numbers of low-scoring students. For example, a recent study in a large Texas city found that student dropouts and push outs accounted for most of the gains in high school student test scores, especially for minority students. The introduction of a high-stakes test linked to school ratings in the 10th grade led to sharp increases in 9th grade student retention and student dropout and disappearance. Of the large share of students held back in the 9th grade, most of them African American and Latino, only 12% ever took the 10th grade test that drove school rewards. Schools that retained more students at grade 9 and lost more through dropouts and disappearances boosted their accountability ratings the most. Overall, fewer than half of all students who started 9th grade graduated within 5 years, even as test scores soared.[10]

Paradoxically, NCLB's requirement for disaggregating data and tracking progress for each subgroup of students increases the incentives for eliminating those at the bottom of each subgroup, especially where schools have little capacity to improve the quality of services such students receive. Table 1 shows how this can happen. At "King Middle School," average scores increased from the 70th to the 72nd percentile

between the 2002 and 2003 school year, and the proportion of students in attendance who met the proficiency standard (a score of 65) increased from 66% to 80%—the kind of performance that a test-based accountability system would reward. Looking at subgroup performance, the proportion of Latino students meeting the standard increased from 33% to 50%, a steep increase.

However, not a single student at King improved his or her score between 2002 and 2003. In fact, the scores of every single student in the school went down over the course of the year. How could these steep improvements in the school's average scores and proficiency rates have occurred? A close look at Table 1 shows that the major change between the two years was that the lowest-scoring student, Raul, disappeared. As has occurred in many states with high stakes-testing programs, students who do poorly on the tests—special needs students, new English language learners, those with poor attendance, health, or family problems—are increasingly likely to be excluded by being counseled out, transferred, expelled, or by dropping out.

TABLE 1.—KING MIDDLE SCHOOL: REWARDS OR SANCTIONS?

The Relationship between Test Score Trends and Student Populations

	2002–03	2003–04
Laura	100	90
James	90	80
Felipe	80	70
Kisha	70	65
Jose	60	55
Raul	20	
	Ave. Score = 70 % meeting standard = 66%	Ave. Score = 72 % meeting standard = 80%

This kind of result is not limited to education. When one state decided to rank cardiac surgeons based on their mortality rates, a follow up investigation found that surgeons' ratings went up as they stopped taking on high-risk clients. These patients were referred out of state if they were wealthy, or were not served, if they were poor.

The three national professional organizations of measurement experts have called attention to such problems in their joint Standards for Educational and Psychological Testing, which note that:

Beyond any intended policy goals, it is important to consider potential unintended effects that may result from large-scale testing programs. Concerns have been raised, for instance, about narrowing the curriculum to focus only on the objectives tested, restricting the range of instructional approaches to correspond to the testing format, increasing the number of dropouts among students who do not pass the test, and encouraging other instructional or administrative practices that may raise test scores without affecting the quality of education. It is important for those who mandate tests to consider and monitor their consequences and to identify and minimize the potential of negative consequences.[11]

Professional testing standards emphasize that no test is sufficiently reliable and valid to be the sole source of important decisions about student placements, promotions, or graduation, but that such decisions should be made on the basis of several different kinds of evidence about student learning and performance in the classroom. For example, Standard 13.7 states:

In educational settings, a decision or characterization that will have major impact on a student should not be made on the basis of a single test score. Other relevant information should be taken into account if it will enhance the overall validity of the decision.[12]

The Psychological Standards for Testing describe several kinds of information that should be considered in making judgments about what a student knows and can do, including alternative assessments that provide other information about performance and evidence from samples of school work and other aspects of the school record, such as grades and classroom observations. These are particularly important for students for whom traditional assessments are not generally valid, such as English language learners and special education students. Similarly, when evaluating schools, it is important to include measures of student progress through school, coursework and grades, and graduation, as part of the record about school accomplishments.

Evaluating Learning Well

Indicators beyond a single test score are important not only for reasons of validity and fairness in making decisions, but also to assess important skills that most standardized tests do not measure. Current accountability reforms are based on the idea that standards can serve as a catalyst for states to be explicit about learning goals, and the act of measuring progress toward meeting these standards is an important force toward developing high levels of achievement for all students. However, an on-demand test taken in a limited period of time on a single day cannot measure all that is important for students to know and be able to do. A credible accountability system must rest on assessments that are balanced and comprehensive with respect to state standards. Multiple-choice and short-answer tests that are currently used to measure standards in many states do not adequately measure the complex thinking, communication, and problem solving skills that are represented in national and state content standards.

Research on high-stakes accountability systems shows that, "what is tested is what is taught," and those standards that are not represented on the high stakes assessment tend to be given short shrift in the curriculum.[13] Students are less likely to engage in extended research, writing, complex problem-solving, and experimentation when the accountability system emphasizes short-answer responses to formulaic problems. These higher order thinking skills are those very skills that often are cited as essential to maintaining America's competitive edge and necessary for succeeding on the job, in college, and in life. As described by Achieve, a national organization of governors, business leaders, and education leaders, the problem with measures of traditional on-demand tests is that they cannot measure many of the skills that matter most for success in the worlds of work and higher education:

States * * * will need to move beyond large-scale assessments because, as critical as they are, they cannot measure everything that matters in a young person's education. The ability to make effective oral arguments and conduct significant research projects are considered essential skills by both employers and postsecondary educators, but these skills are very difficult to assess on a paper-and pencil test.[14]

One of the reasons that U.S. students fall further and further behind their international counterparts as they go through school is because of differences in curriculum and assessment systems. International studies have found that the U.S. curriculum focuses more on superficial coverage of too many topics, without the kinds of in-depth study, research, and writing needed to secure deep understanding. To focus on understanding, the assessment systems used in most high-achieving countries around the world emphasize essay questions, research projects, scientific experiments, oral exhibitions and performances that encourage students to master complex skills as they apply them in practice, rather than multiple-choice tests.

As indicators of the growing distance between what our education system emphasizes and what leading countries are accomplishing educationally, the U.S. currently ranks 28th of 40 countries in the world in math achievement—right above Latvia—and 19th of 40 in reading achievement on the international PISA tests that measure higher-order thinking skills. And while the top-scoring nations—including previously low-achievers like Finland and South Korea—now graduate more than 95% of their students from high school, the U.S. is graduating about 75%, a figure that has been stagnant for a quarter century and, according to a recent ETS study, is now declining. The U.S. has also dropped from 1st in the world in higher education participation to 13th, as other countries invest more resources in their children's futures.

Most high-achieving nations' examination systems include multiple samples of student learning at the local level as well as the state or national level. Students' scores are a composite of their performance on examinations they take in different content areas—featuring primarily open-ended items that require written responses and problem solutions—plus their work on a set of classroom tasks scored by their teachers according to a common set of standards. These tasks require them to conduct apply knowledge to a range of tasks that represent what they need to be able to do in different fields: find and analyze information, solve multi-step real-world problems in mathematics, develop computer models, demonstrate practical applications of science methods, design and conduct investigations and evaluate their results, and present and defend their ideas in a variety of ways. Teaching to these assessments prepares students for the real expectations of college and of highly skilled work.

These assessments are not used to rank or punish schools, or to deny promotion or diplomas to students. In fact, several countries have explicit proscriptions against such practices. They are used to evaluate curriculum and guide investments in professional learning—in short, to help schools improve. By asking students to show what they know through real-world applications of knowledge, these nations' assess-

ment systems encourage serious intellectual activities on a regular basis. The systems not only measure important learning, they help teachers learn how to design curriculum and instruction to accomplish this learning.

It is worth noting that a number of states in the U.S. have developed similar systems that combine evidence from state and local standards-based assessments to ensure that multiple indicators of learning are used to make decisions about individual students and, sometimes, schools. These include Connecticut, Kentucky, Maine, Nebraska, New Hampshire, Oregon, Rhode Island, Pennsylvania, Vermont, and Wyoming, among others. However, many of these elements of state systems are not currently allowed to be used to gauge school progress under NCLB.

Encouraging these kinds of practices could help improve learning and guide schools toward more productive instruction. Studies have found that performance assessments that are administered and scored locally help teachers better understand students' strengths, needs, and approaches to learning, as well as how to meet state standards.[15] Teachers who have been involved in developing and scoring performance assessments with other colleagues have reported that the experience was extremely valuable in informing their practice. They report changes in both the curriculum and their instruction as a result of thinking through with colleagues what good student performance looks like and how to better support student learning on specific kinds of tasks.

These goals are not well served by external testing programs that send secret, secured tests into the school and whisk them out again for machine scoring that produces numerical quotients many months later. Local performance assessments provide teachers with much more useful classroom information as they engage teachers in evaluating how and what students know and can do in authentic situations. These kinds of assessment strategies create the possibility that teachers will not only teach more challenging performance skills but that they will also be able to use the resulting information about student learning to modify their teaching to meet the needs of individual students. Schools and districts can use these kinds of assessments to develop shared expectations and create an engine for school improvement around student work.

Research on the strong gains in achievement shown in Connecticut, Kentucky, and Vermont in the 1990s attributed these gains in substantial part to these states' performance-based assessment systems, which include such local components, and related investments in teaching quality.[16] Other studies in states like California, Maine, Maryland, and Washington,[17] found that teachers assigned more ambitious writing and mathematical problem solving, and student performance improved, when assessments included extended writing and mathematics portfolios and performance tasks. Encouraging these kinds of measures of student performance is critical to getting the kind of learning we need in schools.

Not incidentally, more authentic measures of learning that go beyond on-demand standardized tests to look directly at performance are especially needed to gain accurate measures of achievement for English language learners and special needs students for whom traditional tests are least likely to provide valid measures of understanding.[18]

What Indicators Might be Used to Gauge School Progress?

A key issue is what measures should be used to determine Adequate Yearly Progress (AYP) or the alternative tools that are used for addressing NCLB's primary goals, e.g. assuring high expectations for all students, and helping schools address the needs of all students. Current AYP measures are too narrow in several respects: They are based exclusively on tests which are often not sufficient measures of our educational goals; they ignore other equally important student outcomes, including staying in school and engaging in rigorous coursework; they ignore the growth made by students who are moving toward but not yet at a proficiency benchmark, as well as the gains made by students who have already passed the proficiency benchmark; and they do not provide information or motivation to help schools, districts, and states improve critical learning conditions.

This analysis suggests that school progress should be evaluated on multiple measures of student learning—including local and state performance assessments that provide evidence about what students can actually do with their knowledge—and on indicators of other student outcomes, including such factors as student progress and continuation through school, graduation, and success in rigorous courses. The importance of these indicators is to encourage schools to keep students in school and provide them with high-quality learning opportunities—elements that will improve educational opportunities and attainment, not just average test scores.

To these two categories of indicators, I would add indicators of learning conditions that point attention to both learning opportunities available to students (e.g. rig-

orous courses, well-qualified teachers) and to how well the school operates. In the business world, these kinds of measures are called leading indicators, which represent those things that employees can control and improve upon. These typically include evidence of customer satisfaction, such as survey data, complaints and repeat orders; as well as of employee satisfaction and productivity, such as employee turnover, project delays, evidence of quality and efficiency in getting work done; reports of work conditions and supports, and evidence of product quality.

Educational versions of these kinds of indicators are available in many state accountability systems. For example, State Superintendent Peter McWalters noted in his testimony to this committee that Rhode Island uses several means to measure school learning conditions. Among them is an annual survey to all students, teachers, and parents that provides data on "Learning Support Indicators" measuring school climate, instructional practices, and parental involvement. In addition, Rhode Island, like many other states, conducts visits to review every school in the state every five years, not unlike the Inspectorate system that is used in many other countries. These kinds of reviews can examine teaching practices, the availability and equitable allocation of school resources, and the quality of the curriculum, as it is enacted.

Ideally, evaluation of school progress would be based on a combination of these three kinds of measures and would emphasize gains and improvement over time, both for the individual students in the school and for the school as a whole. Along with data about student characteristics, an indicator system could include:

- Measures of student learning: both state tests and local assessments, including performance measures that assess higher-order thinking skills and understanding, including student work samples, projects, exhibitions, or portfolios.
- Measures of additional student outcomes: data about attendance, student grade-to-grade progress (promotion / retention rates) and continuation through school (ongoing enrollment), graduation, and course success (e.g. students enrolled in, passing, and completing rigorous courses of study).
- Measures of learning conditions, data about school capacity, such as teacher and other staff quality, availability of learning materials, school climate (gauged by students', parents', and teachers' responses to surveys), instructional practices, teacher development, and parental engagement.

These elements should be considered in the context of student data, including information about student mobility, health, and welfare (poverty, homelessness, foster care, health care), as well as language background, race / ethnicity, and special learning needs—not a basis for accepting differential effort or outcomes, but as a basis for providing information needed to interpret and improve schools' operations and outcomes.

How Might Indicators be Used to Determine School Progress and Improvement Strategies?

The rationale for these multiple indicators is to build a more powerful engine for educational improvement by understanding what is really going on with students and focusing on the elements of the system that need to change if learning is to improve. High-performing systems need a regular flow of useful information to evaluate and modify what they are doing to produce stronger results. State and local officials need a range of data to understand what is happening in schools and what they should do to improve outcomes. Many problems in local schools are constructed or constrained by district and state decisions that need to be highlighted along with school-level concerns. Similarly, at the school level, teachers and leaders need information about how they are doing and how their students are doing, based in part on high-quality local assessments that provide rich, timely insights about student performance.

Some states and districts have successfully put some of these indicators in place. The federal government could play a leadership role by not only encouraging multiple measures for assessing school progress and conditions for learning but by providing supports for states to build comprehensive databases to track these indicators over time, and to support valid, comprehensive information systems at all levels.[19]

If we think comprehensively about the approach to evaluation that would encourage fundamental improvements in schools, several goals emerge. First, determinations of school progress should reflect an analysis of schools' performance and progress along several key dimensions. Student learning should be evaluated using multiple measures that provide comprehensive and valid information for all subpopulations. Targets should be based on sensible goals for student learning, examining growth from where students start, setting growth targets in relation to that starting point, and pegging "proficiency" at a level that represents a challenging but realistic standard, perhaps at the median of current state proficiency standards.

Targets should also ensure appropriate assessment for special education students and English language learners and credit for the gains these students make over time. And analysis of learning conditions including the availability of materials, facilities, curriculum opportunities, teaching, and leadership should accompany assessments of student learning.

A number of states already have developed comprehensive indicator systems that can be sources of such data, and the federal government should encourage states to propose different means for how to aggregate and combine these data. In addition, many states' existing assessment systems already provide different ways to score and combine state reference tests with local testing systems, locally administered performance tasks (which are often scored using state standards), and portfolios.[20]

For evaluating annual progress, one likely approach would be to use an index of indicators, such as California's Academic Performance Index, which can include a weighted combination of data about state and local tests and assessments as well as other student outcome indicators like attendance, graduation, promotion rates, participation and pass rates or grades for academic courses. Assessment data from multiple sources and evidence of student progression through / graduation from school would be required components. Key conditions of learning, such as teacher qualifications, might also be required. Other specific indicators might be left to states, along with the decision of how much weight to give each component, perhaps within certain parameters (for example, that at least 50 percent of a weighted index would reflect the results of assessment data).

Within this index, disaggregated data by race/ethnicity and income could be monitored on the index score, or on components of the overall index, so that they system pays ongoing attention to progress for groups of students. Wherever possible these measures should look at progress of a constant cohort of students from year to year, so that actual gains are observed, rather than changes in averages due to changes in the composition of the student population. Furthermore, gains for English language learners and special education students should be evaluated on a growth model that ensures appropriate testing based on professional standards and measures individual student growth in relation to student starting points.

Non-academic measures such as improved learning climate (as measured by standard surveys, for example, to allow trend analysis over time), instructional capacity (indicators regarding the quality of curriculum, teaching, and leadership), resources, and other contributors to learning could be included in a separate index on Learning Conditions, on which progress is also evaluated annually as part of both school, district, and state assessment.

Once school progress indicators are available, a judgment must be made about whether a school has made adequate progress on the index or set of indicators. If the law is to focus on supporting improvement it will be important to look at continuous progress for all students in a school rather than the "status model" that has been used in the past. A progress model would recognize the reasonable success of schools that deserve it. Rather than identifying a school as requiring intervention when a single target is missed (for example, if 94% of economically disadvantaged students take the mathematics test one year instead of 95%), a progress model would gauge whether the overall index score increases, with the proviso that the progress of key subgroups continues to be examined, with lack of progress a flag for intervention.

The additional use of the indicators schools and districts have assembled would be in the determination of what kind of action is needed if a school does not make sufficient progress in a year. To use resources wisely, the law should establish a graduated system of classification for schools and districts based on their rate of progress, ranging from state review to corrective actions to eventual reconstitution if such efforts fail over a period of time. States should identify schools and districts as requiring intervention based both on information about the overall extent of progress from the prior year(s) and on information about specific measures in the system of indicators—for example, how many progress indicators have lagged for how long. This additional scrutiny would involve a school review by an expert team—much like the inspectorate systems in other countries—that conducts an inspection of the school or LEA and analyzes a range of data, including evidence of individual and collective student growth or progress on multiple measures; analysis of student needs, mobility, and population changes; and evaluation of school practices and conditions. Based on the findings of this review, a determination would be made about the nature of the problem and the type of school improvement plan needed. The law should include the explicit expectation that state and district investments in ensuring adequate conditions for learning must be part of this plan.

The overarching goal of the ESEA should be to improve the quality of education students receive, especially those traditionally least well served by the current system. To accomplish this, the measures used to gauge school progress must motivate continuous improvement and attend to the range of school outcomes and conditions that are needed to ensure that all students are educated to higher levels.

ENDNOTES

[1] See, e.g. L. Darling-Hammond, No Child Left Behind and High School Reform, Harvard Education Review, 76, 4 (Winter 2006), pp. 642-667. *http://www.edreview.org/harvard06/2006/wi06/w06darli.htm*

L. Darling-Hammond, From 'Separate but Equal' to 'No Child Left Behind': The Collision of New Standards and Old Inequalities. In Deborah Meier and George Wood (eds.), Many Children Left Behind, pp. 3-32. NY: Beacon Press, 2004.

[2] Linda Darling-Hammond, Elle Rustique-Forrester, & Raymond Pecheone (2005). Multiple measures approaches to high school graduation: A review of state student assessment policies. Stanford, CA: Stanford University, School Redesign Network.

[3] Allington, R. L. & McGill-Franzen, A. (1992). Unintended effects of educational reform in New York, Educational Policy, 6 (4): 397-414; Figlio, D.N. & Getzler, L.S. (2002, April). Accountability, ability, and disability: Gaming the system? National Bureau of Economic Research.

[4] W. Haney (2000). The myth of the Texas miracle in education. Education Policy Analysis Archives, 8 (41): Retrieved Jan. 24, 08 from: *http://epaa.asu.edu/epaa/v8n41/*

[5] Smith, F., et al. (1986). High school admission and the improvement of schooling. NY: New York City Board of Education; Darling-Hammond, L. (1991). The Implications of Testing Policy for Quality and Equality, Phi Delta Kappan, November 1991: 220-225; Heilig, J. V. (2005), An analysis of accountability system outcomes. Stanford University.

[6] For recent studies examining the increases in dropout rates associated with high-stakes testing systems, see Advocates for Children (2002). Pushing out at-risk students: An analysis of high school discharge figures—a joint report by AFC and the Public Advocate. *http://www.advocatesforchildren.org/pubs/pushout-11-20-02.html;* W. Haney (2002). Lake Wobegone guaranteed: Misuse of test scores in Massachusetts, Part 1. Education Policy Analysis Archives, 10(24). *http://epaa.asu.edu/epaa/v10n24/;* J. Heubert & R. Hauser (eds.) (1999). High stakes: Testing for tracking, promotion, and graduation. A report of the National Research Council. Washington, D.C.: National Academy Press; B.A. Jacob (2001). Getting tough? The impact of high school graduation exams. Education and Evaluation and Policy Analysis 23 (2): 99-122; D. Lilliard, & P. DeCicca (2001). Higher standards, more dropouts? Evidence within and across time. Economics of Education Review, 20(5): 459-73;G. Orfield, D. Losen, J. Wald, & C.B. Swanson (2004). Losing our future: How minority youth are being left behind by the graduation rate crisis. Retrieved January 24, 2008 from: *http://www.urban.org/url.cfm?ID=410936; M. Roderick, A.S. Bryk, B.A. Jacob, J.Q. Easton, & E. Allensworth (1999). Ending social promotion: Results from the first two years. Chicago: Consortium on Chicago School Research; R. Rumberger & K. Larson (1998). Student mobility and the increased risk of high school dropout. American Journal of Education, 107: 1-35; E. Rustique-Forrester (in press). Accountability and the pressures to exclude: A cautionary tale from England. Education Policy Analysis Archives; A. Wheelock (2003). School awards programs and accountability in Massachusetts.*

[7] Advocates for Children (2002), Pushing out at-risk students; Heilig (2005), An analysis of accountability system outcomes; Wheelock (2003), School awards programs and accountability.

[8] Heilig, 2005.

[9] Wheelock, 2003

[10] Heilig, 2005.

[11] American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, Standards for Educational and Psychological Testing, Washington DC: American Educational Research Association, 1999, p.142.

[12] AERA, APA, NCME, Standards for Educational and Psychological Testing., p.146.

[13] See for example, Haney (2000). The myth of the Texas miracle; J.L. Herman & S. Golan (1993). Effects of standardized testing on teaching and schools. Educational Measurement: Issues and Practice, 12(4): 20-25, 41-42; B.D. Jones & R. J. Egley (2004). Voices from the frontlines: Teachers' perceptions of high-stakes testing. Education Policy Analysis Archives, 12 (39). Retrieved August 10, 2004 from *http://epaa.asu.edu/epaa/v12n39/;* M.G. Jones, B.D. Jones, B. Hardin, L. Chapman, & T. Yarbrough (1999). The impact of high-stakes testing on teachers and students in North Carolina. Phi Delta Kappan, 81(3): 199-203; Klein, S.P., Hamilton, L.S., McCaffrey, D.F., & Stetcher, B.M. (2000). What do test scores in Texas tell us? Santa Monica: The RAND Corporation; D. Koretz & S. I. Barron (1998). The validity of gains on the Kentucky Instructional Results Information System (KIRIS). Santa Monica, CA: RAND, MR-1014-EDU; D. Koretz, R.L. Linn, S.B. Dunbar, & L.A. Shepard (1991, April). The effects of high-stakes testing: Preliminary evidence about generalization across tests, in R. L. Linn (chair), The Effects of high stakes testing. Symposium presented at the annual meeting of the American Educational Research Association and the National Council on Measurement in Education, Chicago; R.L. Linn (2000). Assessments and accountability. Educational Researcher, 29 (2), 4-16; R.L. Linn, M.E. Graue, & N.M. Sanders (1990). Comparing state and district test results to national norms: The validity of claims that "everyone is above average." Educational Measurement: Issues and Practice, 9, 5-14; W. J. Popham (1999). Why Standardized Test Scores Don't Measure Educational Quality. Educational Leadership, 56(6): 8-15; M.L. Smith (2001). Put to the test: The effects of external testing on teachers. Educational Researcher, 20(5): 8-11.

[14] Achieve, Do graduation tests measure up? A closer look at state high school exit exams. Executive summary. Washington, DC: Achieve, Inc.

[15] L. Darling-Hammond & J. Ancess (1994). Authentic assessment and school development. NY: National Center for Restructuring Education, Schools, and Teaching, Teachers College, Columbia University; B. Falk & S. Ort (1998, September). Sitting down to score: Teacher learning through assessment. Phi Delta Kappan, 80(1): 59-64. G.L. Goldberg & B.S. Rosewell (2000). From perception to practice: The impact of teachers' scoring experience on the performance based instruction and classroom practice. Educational Assessment, 6: 257-290; R. Murnane & F. Levy (1996). Teaching the new basic skills. NY: The Free Press.

[16] J.B. Baron (1999). Exploring high and improving reading achievement in Connecticut. Washington: National Educational Goals Panel. Murnane & Levy (1996); B.M. Stecher, S. Barron, T. Kaganoff, & J. Goodwin (1998). The effects of standards-based assessment on classroom practices: Results of the 1996-97 RAND survey of Kentucky teachers of mathematics and writing. CSE Technical Report. Los Angeles: UCLA National Center for Research on Evaluation, Standards, and Student Testing; S. Wilson, L. Darling-Hammond, & B. Berry (2001). A case of successful teaching policy: Connecticut's long-term efforts to improve teaching and learning. Seattle: Center for the Study of Teaching and Policy, University of Washington.

[17] C. Chapman (1991, June). What have we learned from writing assessment that can be applied to performance assessment?. Presentation at ECS/CDE Alternative Assessment Conference, Breckenbridge, CO; J.L.Herman, D.C. Klein, T.M. Heath, S.T. Wakai (1995). A first look: Are claims for alternative assessment holding up? CSE Technical Report. Los Angeles: UCLA National Center for Research on Evaluation, Standards, and Student Testing; D. Koretz, K., J. Mitchell, S.I. Barron, & S. Keith (1996). Final Report: Perceived effects of the Maryland school performance assessment program CSE Technical Report. Los Angeles: UCLA National Center for Research on Evaluation, Standards, and Student Testing; W.A. Firestone, D. Mayrowetz, & J. Fairman (1998, Summer). Performance-based assessment and instructional change: The effects of testing in Maine and Maryland. Educational Evaluation and Policy Analysis, 20: 95-113; S. Lane, C.A. Stone, C.S. Parke, M.A. Hansen, & T.L. Cerrillo (2000, April). Consequential evidence for MSPAP from the teacher, principal and student perspective. Paper presented at the annual meeting of the National Council on Measurement in Education, New Orleans, LA; B. Stecher, S. Baron, T. Chun, T., & K. Ross (2000) The effects of the Washington state education reform on schools and classroom. CSE Technical Report. Los Angeles: UCLA National Center for Research on Evaluation, Standards, and Student Testing.

[18] Darling-Hammond, Rustique-Forrester, and Pecheone, Multiple Measures.

[19] M. Smith paper (2007). Standards-based education reform: What we've learned, where we need to go. Consortium for Policy Research in Education.

[20] At least 27 states consider student academic records, coursework, portfolios of student work, and performance assessments, like research papers, scientific experiments, essays, and senior projects in making the graduation decision. Darling-Hammond, Rustique-Forrester, and Pecheone, Multiple Measures.

[The prepared statement of the Coalition Promoting School Success for All Children, submitted by Mr. Hare, follows:]

Prepared Statement of the Coalition Promoting School Success for All Children

Schools today are facing two significant challenges: to improve academic achievement, including meeting the requirements of the No Child Left Behind Act (NCLB), and to foster a school climate that promotes learning and that will reduce discipline problems. These issues are linked. Research shows that academic, social and behavioral problems are so connected that interventions targeting one often affect the others. Consequently, effective schools support and foster both high academic and high behavioral standards.

Integrated efforts to address academic and non-academic barriers to learning are more likely to help schools meet academic requirements than efforts that focus on academics alone. Amendments to the No Child Left Behind Act that encourage schools to improve the climate for learning, such as by adopting school-wide positive behavior supports (PBS) will help schools to address both students' social and emotional needs and lead to improvements in academic outcomes.

Both educators and the general public cite disciplinary issues (including a perceived increase in drugs, violence, gangs and weapons) as the number one problem in schools, and teachers say they feel unprepared to manage problem behaviors. In efforts to improve school safety, many school officials react by taking a tough approach using suspension or expulsion extensively or trying to remove persistent troublemakers from school. But research shows that rigid and inflexible approaches to discipline do not work and, further, that they disproportionately harm students of color and students with disabilities. In contrast, positive and relational approaches can improve the school environment without resorting to exclusionary practices.

To address discipline problems, help children learn and meet today's demands for high academic standards, schools should improve the learning environment for all students and provide the supports needed by those with more significant problems. One effective approach now being adopted by education systems around the country

is school-wide PBS. PBS improves student behavior by reinforcing desired behavior and eliminating inadvertent reinforcements for problem behavior. This requires understanding the reason for the behavior and addressing the underlying cause.

Schools implementing school-wide PBS programs can experience anywhere from a 20-60 percent reduction in disciplinary problems, as well as improved social climate and academic performance. Research shows increases in both reading comprehension and math test scores on standardized tests in schools implementing PBS well. In addition to improved academic outcomes, school-wide PBS improves other variables related to student success, including increased student attendance, fewer expulsions and suspensions, increased classroom instructional time and academic engagement. There is more time for student instruction and a reduction in hours spent by teachers and administrators addressing problem behavior. With effective academics and behavior intertwined, school-wide PBS is instrumental in changing school climate and allowing for more effective instruction.

School-wide PBS is not a program, but a system, based on decades of behavioral and biomedical research. Using an approach adapted from the public health field, PBS uses a three-tiered system of prevention and support that addresses the spectrum of behavioral needs and serves all children—from those with behavior issues that are typical of their developmental stage, as well as those at risk for or already exhibiting challenging behaviors. For the general student body—roughly 80 percent of students—universal implementation of PBS in the school (known as Tier One) may be sufficient. Tier One focuses on changing environmental stimuli that contribute to disruptive behavior and on changing adult behavior in school so that all staff are supportive and consistently teach, reinforce, and model expected behaviors.

Five to 15 percent of students who do not respond to universal methods do respond to more specialized attention, an example being a group intervention. These children fall into PBS Tier Two. An intervention in Tier Two might involve group sessions where students problem-solve and come up with strategies to prevent the problem behavior.

Finally, some children with the most challenging behavior need individualized services (Tier Three). Often these are children with serious emotional disorders and extreme functional impairment. They represent three to seven percent of all school-age children. These students should be involved in a comprehensive home, school and community plan using individualized services and techniques coordinated across agencies.

Training is an integral part of a PBS initiative and must be continual as new staff are hired or as additional schools in the district adopt PBS. PBS uses in-school coaches to help translate training into practice and to support staff who are implementing school-wide PBS. These in-school coaches are often school psychologists, social workers or counselors who must themselves be trained in school-wide PBS. In addition, external coaches are employed, who typically work with a number of schools, collaborating with the in-school coach and providing feedback to the state and region/district, as well as providing guidance to individual schools.

It is important for best practices, including particularly family involvement and social-emotional learning, must be followed. School-wide PBS provides a strong platform for related programming, such as specific social-emotional learning programs and other youth-development practices. It also requires family buy-in, participation and support. Families play a crucial role in implementation of school-wide PBS.

Two examples of statewide PBS initiatives are Illinois and Colorado. These initiatives are well-established and continue to grow and reach more schools throughout each state.

- The state of Illinois is a pioneer in creating a statewide comprehensive PBS initiative. In eight years, Illinois has built and sustained the Illinois PBIS Network, providing assistance in implementation of PBS to 600 schools, or around 14 percent of all public schools in the state. Research in Illinois shows that implementation of school-wide PBS is linked to improved perception of school safety, as well as an improved proportion of third graders meeting state reading standards.
- The Colorado School-Wide PBS Initiative is a joint venture between the Prevention Initiatives and Exceptional Student Services Unit in the Colorado Department of Education. The initiative began providing support and assistance to schools implementing PBS in the 2002/2003 school year. They are now in 405 schools in the state, operating in eight regions and 48 school districts. They continue to work toward their goal of establishing and maintaining effective school environments maximizing academic achievement and behavioral competence for all learners in Colorado.

Although many of the early school-wide PBS initiatives were initiated under the special education umbrella, the approach is more effective when implemented in a systematic, school-wide effort. PBS is not just helpful for students in special edu-

cation or those with disabilities. Instead, it is a means for all students to increase academic achievement, improve social behavior and learn self-management.

School-wide PBS complements the academic standards articulated in NCLB. The following amendments to the NCLB Act would create avenues for schools to more easily implement school-wide PBS.

• Encourage states to use Title I funds to support school-wide PBS

The Individuals with Disabilities Education Act (IDEA) promotes Early Intervening Services by allowing local education agencies (LEAs) to use up to 15 percent of IDEA Part B funds to implement services for students not identified as needing special education, but who do need additional academic and behavioral support to succeed in the general education environment. Such funds have been used by a number of districts to support school-wide PBS. These IDEA funds may be used in coordination with funds from ESEA to carry out these activities and services.

However, there is no comparable language in the ESEA to allow Title I funds to be used for Early Intervening Services. Authorization to use ESEA funds for Early Intervening Services, including specifically school-wide PBS, would facilitate expansion of these extremely promising approaches. Since this is not a mandate, LEAs would have flexibility as to whether or not to use the funds for these purposes.

Section 1003 of NCLB authorizes states to reserve up to two to four percent of Title I funds for school improvement to meet state responsibilities under Section 1116 and 1117. Amending Sections 1003, 1116 and 1117 to include implementation of school-wide PBS will permit states to use their Title I state reserve funds to promote and support such initiatives.

Encourage LEAs to address school climate issues

Schools must promote an environment that is safe and conducive to learning, providing the foundation on which other programming and support can be built. School-wide PBS contributes to attitudinal change, creating a culture where there is a shared sense of responsibility. The positive school climate that comes from a school implementing school-wide PBS promotes learning by reducing discipline problems and addressing the social and emotional development of students.

Section 1114 authorizes LEAs to use funds for school-wide programs. Adding school-wide PBS to this section increases the likelihood that Title I funds will be used for this purpose.

Establish an Office of Specialized Instructional Support Personnel

To raise the visibility of critical issues and to coordinate across the various departmental agencies, an Office of Specialized Instructional Support Personnel should be established in the Office of the Deputy Secretary in the Department of Education. The purpose of the Office should be to administer, coordinate and carry out programs and activities concerned with providing specialized instructional support services in schools, delivered by trained, qualified specialized instructional support personnel.

Activities governed by such an office would include:

• Improving academic achievement and educational results for students through improved instructional support services in schools, including provision of early intervening services to general education students and of the related services required under IDEA;

• Administering, coordinating and carrying out programs and activities concerned with providing specialized instructional support services in schools;

• Promoting a trained, qualified specialized instructional support workforce in schools; and

• Providing technical assistance to local and state education agencies in provision of effective, scientifically-based specialized instructional support services.

Amend the Safe & Drug Free Schools and Communities Program

The Safe & Drug Free Schools and Communities Program should be amended to incorporate the concept of creating a safe and effective school climate that is conducive to learning. Specifically, this program should be amended by:

• Changing the title to expand the purpose of the program to address school climate in general. The amended title would read: Creating a Safe and Effective School Climate Conducive for Learning. The short title would read: The Safe and Effective Schools Act.

• Clarifying that the purpose of Part A is to address whole school climate and to prevent violence and illegal use of alcohol, tobacco and drugs.

• Adding to paragraph (1), the list of purposes for which states may use funds under this part, grants to states or LEAs and consortia of such agencies for the establishment, operation and improvement of local programs relating to improving the

school-wide climate (including the implementation of school-wide PBS and other programs).

The Coalition Promoting School Success for All Children commend you for your ongoing leadership and commitment to education and ensuring no children are left behind, including those with mental health needs. We look forward to working with you to strengthen and protect the provisions in No Child Left Behind by including school-wide positive behavior supports that will be beneficial to all students.

The Coalition Promoting School Success for All Children includes the following organizations:

Advocacy Institute
American Counseling Association
American School Counselor Association
American Occupational Therapy Association
Bazelon Center for Mental Health Law
Children & Adults with Attention-Deficit/Hyperactivity Disorder
Council for Children with Behavioral Disorders
Council for Exceptional Children
Easter Seals
Federation of Families for Children's Mental Health
Georgetown University Center for Child and Human Development
Mental Health America
National Alliance on Mental Illness
National Association of School Psychologists
National Association of State Directors of Special Education
National Association of State Mental Health Program Directors
National Down Syndrome Congress
School Social Work Association of America

The Bazelon Center for Mental Health Law chairs the Coalition. Please contact Laurel Stine for any additional information.

Chairman KILDEE. I would now like to introduce the very distinguished panel of witnesses here with us this afternoon. John "Jack" Jennings is President of the Center on Education Policy in Washington, D.C., the center promotes school improvement through various publications, by assisting States and school districts, and by arranging national meetings and conferences. From 1967 through 1994, as a staff member and Chief Counsel of this committee, Mr. Jennings was deeply involved in virtually every major national education debate. I learned the word "disaggregated data" from Mr. Jennings at that time.

Mr. Rick Melmer is South Dakota Secretary of Education and President-Elect of the Council of Chief State School Officers. Previously he served as school superintendent in Watertown, South Dakota and Sioux City, Iowa. Dr. Melmer has said that his role as Secretary is to shine a spotlight on the future of education in South Dakota.

And the Honorable Kathleen Straus, who I have known forever, is President of the Michigan State Board of Education. She has dedicated her career to improving public education and community services in Michigan. Prior to serving on the State board, Ms. Straus was President of the Center for Creative Studies, a nationally recognized arts education institution in Detroit. She also has worked for the Michigan Association of School Boards and the Michigan Senate Education Committee.

Dr. Carol Johnson has been Superintendent of the Memphis, Tennessee City Schools since 2003 and is a member of the Executive Committee of the Council of the Great City Schools. In 2006, for the first time since No Child Left Behind was passed, the Memphis schools were declared to be in good standing by the State. The Memphis City Schools also have been recognized as a model for

middle school student engagement and is one of the top districts for music education.

And Dr. Chester Finn, whom I have known for a long, long time, and good to see you again, is President of the Thomas B. Fordham Foundation, which supports education reform research and projects. Dr. Finn also is Chairman of the Hoover Institution's Task Force on K-12 Education. Among his other positions he served as Assistant Secretary of Education for Research and Improvement under President Reagan and as a Professor of Education and Public Policy at Vanderbilt University.

And welcome to all our witnesses. For those of you who have not testified before this subcommittee I will explain to you our lighting system and the 5-minute rule. Everyone, including members on the dais up here, is limited to 5 minutes of presentation or questioning. The green light will be illuminated when you begin to speak. When you see the yellow light, it means that you have one minute remaining. When you see the red light it means your time has expired and you need to conclude your testimony. But there is no ejection seat there, so I will allow you to finish your thought anyway. Please be certain as you testify to turn on and speak into the microphone in front of you and turn it off when you are finished. We will now hear from our first witness, Mr. Jennings.

STATEMENT OF JACK JENNINGS, PRESIDENT, CENTER ON EDUCATION POLICY; ACCOMPANIED BY DIANE STARK RENTNER, DIRECTOR OF NATIONAL PROGRAMS AT CEP; NANCY KOBER, CEP'S PRINCIPAL CONSULTANT; AND SUNNY BECKER, PRINCIPAL STAFF SCIENTIST OF THE HUMAN RESOURCES RESEARCH ORGANIZATION

Mr. JENNINGS. Thank you, Chairman Kildee and Congressman Castle and other members of the committee. I spent nearly half my life in this room or around this room or with the committee, and so I feel like I am coming home. They say you can never come home, but at least for this afternoon I will be home, and I think I will be home once I hear the questions coming from all different directions. Then I will know.

I am accompanied by Diane Stark Rentner, who is sitting behind me. She is the Deputy at the Center on Education Policy, and she worked for the committee for 6 years. And then behind me on the other side is Nancy Kober, who worked for the committee for 12 years. And she is the principal consultant for us. And in the middle is Sunny Becker, who is with the Human Resources Company Corporation, HumRRO, that was our contractor on this project. I know you have strict time restraints, so let me just shortly summarize our report.

First of all, the Center on Education Policy is an independent body. We get all our money from charitable foundations. We have no connection with the government, no connection with any other organization, any teachers group, any business group, any other organization. We are totally independent. And we have been reviewing No Child Left Behind since 2002.

But 2 years ago we decided that the key questions involved with No Child Left Behind were two. One was has student achievement in reading and math increased since No Child Left Behind was en-

acted? And secondly, have achievement gaps between different subgroups of students narrowed since No Child Left Behind was enacted? Those are the key questions.

And sometimes too often we spend time on a mechanism such as how many children are getting tutoring, how many children are getting school choice without looking at the fundamental purposes. So we decided to design a study that would answer these questions. And we have five unique features in our study.

First of all, we asked all States to participate. And every State participated and every State verified their data so we know this is correct data.

Secondly, we only use comparable test results from year to year. If a State changed its test we would not use the test results. These have to be comparable results.

Thirdly, we use pre and post-NCLB test scores so that we could go as far back as we could to see the effects of the legislation.

Fourthly, we use two different types of statistical analysis. One is looking at proficiency scores, the other is looking at net effects scores. And one analysis makes up for the deficiencies of the others so that you get a more complete picture.

And lastly, we use consistent rules in looking at the data across all States. Another unique feature was that we ensured that we had a group of experts that was balanced. So we have five experts, several of them have written articles and books against No Child Left Behind, several of them have supported No Child Left Behind. In fact, one of them heads up one of the administration's committees dealing with No Child Left Behind. This group is not only varied in its point of view, but it also has great expertise in terms of assessment, research and policy, and they were of great assistance to us.

The five major conclusions we reached are these. First, in most States with 3 or more years of comparable test data student achievement in reading and math has gone up since 2002.

Secondly, there is more evidence of achievement gaps between groups of students narrowing since 2002 than of gaps widening. Still the magnitude of the gaps is very substantial. It is common to have gaps of 25, 30 percent between different groups of students.

Thirdly, we can only get comparable data from 13 States going back to 1999, but nine of those States showed that they had greater increases in test scores after 2002 rather than before 2002.

Fourthly, and this is a very important point, it is very difficult, if not impossible, to determine to the extent to which these trends in test results have occurred because of No Child Left Behind. Since 2002 States, school districts and schools have simultaneously implemented many different, but interconnected policies to raise achievement. In other words, we cannot show direct causality between increased test scores and No Child Left Behind because all the other things have been going on at the same time.

Fifthly, although NCLB is premised on the idea that there should be public information and there should be test score data available to policy makers, the press and the public so they will understand the effectiveness of education, we found gaps and holes in terms of the information that was available. And it took us a considerable effort in order to obtain this information.

As cooperative as the States were, they are very overburdened, they are very pressed for personnel and time, and it is difficult for them to keep up.

Now, let me mention what we think the possible explanations are for these increased test trends, and these are four possible reasons. And we cannot attribute direct causality in whatever proportion to any of these reasons but these are the logical reasons.

One, in fact students have increased learning. That sounds so obvious but in today's cynical environment you have to say well, students may know more. And that is probably a major factor.

Secondly, there is teaching to the test. Especially in the poorest school districts, the schools with largest percentages of children of color there is teaching to the test.

Thirdly, there are subtle changes made in tests that make it easier to have gains. And lastly there is changes in populations tested.

I will finish by just asking that you consider the data that is required from the States for No Child Left Behind. And we have a recommendation on page 81 of the report. I hope you rip that page out and put it into the law. The public deserves to know test data, they deserve to know as much information as they can about the schools, and you can help to ensure that by enacting those provisions.

Thank you.

[The statement of Mr. Jennings follows:]

Prepared Statement of Jack Jennings, President, Center on Education Policy

Chairman Kildee and Members of the Subcommittee, thank you for the opportunity to appear before you today on trends in student achievement and the No Child Left Behind Act. I am accompanied by Diane Stark Rentner, director of national programs at the Center, Nancy Kober, CEP's principal consultant, and Sunny Becker, principal staff scientist of the Human Resources Research Organization, CEP's contractor for this project.

The Center on Education Policy, a private non-profit organization, is an independent national advocate for effective public schools. We principally accomplish our mission by analyzing policies to determine whether they are in fact helping public schools to become better. A principal focus of our work for the last five years has been the No Child Left Behind Act, since that policy is so significant for public education. Since NCLB was enacted, we have monitored its effects and issued both comprehensive and special reports. Today, I will discuss our latest report which we released this Tuesday at the National Press Club.

Since 2002, NCLB has spurred far-reaching changes in elementary and secondary education, all aimed at accomplishing the same fundamental goal—to improve students' academic achievement. As the Congress prepares to reauthorize the Act, two related questions matter most:

1. Has student achievement in reading and math increased since NCLB was enacted?

2. Have achievement gaps between different subgroups of students narrowed since NCLB was enacted?

To answer these questions, the Center on Education Policy conducted the most comprehensive study of trends in state test scores since NCLB took effect. We carried out this study with advice from a panel of five nationally known experts in educational testing or policy research, and with extensive technical support from HumRRO. Although we collected data from all 50 states, not every state had enough consistent data to do a complete analysis of test score trends in reading and math before and after 2002. Based on the data that states did provide, we reached five main conclusions.

Main Conclusions

1. In most states with three or more years of comparable test data, student achievement in reading and math has gone up since 2002, the year NCLB was enacted.

2. There is more evidence of achievement gaps between groups of students narrowing since 2002 than of gaps widening. Still, the magnitude of the gaps is often substantial.

3. In 9 of the 13 states with sufficient data to determine pre-and post-NCLB trends, average yearly gains in test scores were greater after NCLB took effect than before.

4. It is very difficult, if not impossible, to determine the extent to which these trends in test results have occurred because of NCLB. Since 2002, states, school districts, and schools have simultaneously implemented many different but interconnected policies to raise achievement.

5. Although NCLB emphasizes public reporting of state test data, the data necessary to reach definitive conclusions about achievement were sometimes hard to find or unavailable, or had holes or discrepancies. More attention should be given to issues of the quality and transparency of state test data.

The study that produced these conclusions had several unique features, designed to address the limitations of past research on achievement since 2002. We went to great lengths to gather the most current results on state reading and mathematics tests from all 50 states and to have all states verify the accuracy of their data. Within each state, we limited our analyses to test results that were truly comparable from year to year—in other words, that had not been affected by such factors as the adoption of new tests or changes in the test score students must reach to be considered proficient. We also compared trends before and after 2002 to see whether the pace of improvement has sped up or slowed down since NCLB took effect. We supplemented our analyses of the percentage of students scoring at or above the proficient level—the "magic number" for NCLB accountability—with analyses of effect size, a statistical tool based on average (mean) test scores that addresses some of the problems with the percentage proficient measure. And we analyzed all of the data—which in a typical state included as many as 16,000 individual numbers—as objectively as possible, using a consistent set of rules that were developed without regard to whether they would lead to positive or negative findings.

The rest of this testimony summarizes the findings that led us to the five main conclusions. Further detail can be found in our full report which appears on our Web site, CEP-DC.org.

Gains in Reading and Math Since 2002

To reach national conclusions about reading and math achievement, we first determined the test score trends in each state, looking at both the percentages of students scoring proficient and effect sizes where available. The state trends were then aggregated into a national picture of achievement that included these and other findings:

• The number of states showing gains in test scores since 2002 is far greater than the number showing declines. For example, of the 24 states with percentage proficient and effect size data for middle school reading, 11 demonstrated moderate-to-large gains (average gains of at least 1 percentage point per year) in middle school reading, and only one showed a moderate or larger decline. Five of the 22 states with both percentage proficient and effect size data at the elementary, middle, and high school levels made moderate-to-large gains in reading and math on both measures across all three grade spans. In other words, these five states showed gains according to all of the indicators collected for this study. In reading alone, seven states showed moderate-to-large increases across all three grade spans on both measures. In math alone, nine states showed similar gains across all three grade spans on both measures. The rest of the states had different trends at different grade spans.

• Elementary school math is the area in which the most states showed improvements. Of the 25 states with sufficient data, 22 demonstrated moderateto-large math gains at the elementary level on both the percentage proficient and effect size measures, while none showed moderate or larger declines. Based on percentages proficient alone, 37 of the 41 states with trend data in elementary math demonstrated moderate-to-large gains, while none showed moderate or larger declines.

• More states showed declines in reading and math achievement at the high school level than at the elementary or middle school levels. Still, the number of states with test score gains in high school exceeded the number with declines.

• Analyses of changes in achievement using effect sizes generally produced the same findings as analyses using percentages proficient. But in some cases, the effect

size analysis showed a different trend. In Nevada, for instance, the percentage proficient in high school math decreased, while the average test score increased. In New Jersey the percentage proficient in middle school reading rose slightly, while the average test score dropped.

• When the percentage of students scoring at the proficient level on state tests is compared with the percentage scoring at the basic level on the National Assessment of Educational Progress (NAEP), states show more positive results on their own tests than on NAEP. Moreover, the states with the greatest gains on their own tests were usually not the same states that had the greatest gains on NAEP. The NAEP tests, however, are not aligned with a state's curriculum as state tests are, so NAEP should not be treated as a "gold standard" to invalidate state test results but as an additional source of information about achievement.

Narrowing Achievement Gaps

We analyzed trends in test score gaps for major racial-ethnic subgroups of students, low-income students, students with disabilities, and limited-English-proficient (LEP) students.

We looked at both percentages proficient and effect size data where available; effect size data were harder to come by for subgroups than for students overall. We considered a narrowing or widening of the achievement gap to be a trend for a specific subgroup if it occurred in the same subject (reading or math) across all three grade spans (elementary, middle, and high school). We compiled trends from the 50 states to arrive at these and other national findings:

• Among the states with sufficient data to discern trends by subgroup, the number of states in which gaps in percentages proficient have narrowed since 2002 far exceeds the number of states in which gaps widened.

• For the African-American subgroup, 14 of the 38 states with the necessary data showed evidence that gaps have narrowed in reading across all three grade spans analyzed, while no state had evidence that gaps have widened. In mathematics, 12 states showed these gaps narrowing, while only one state showed the gaps widening. Results were similar for the Hispanic and low-income subgroups.

• As with the percentage proficient, the states in which effect size gaps have narrowed outnumbered the states in which effect size gaps have widened. However, for states with both types of data, there were a number of instances where gap closings in terms of percentages proficient were not confirmed by effect size. Effect sizes seem to give a less rosy picture of achievement gap trends.

• Even for subgroups that showed evidence of gaps narrowing, the gaps in percentages proficient often amounted to 20 percentage points or more, suggesting that it will take a concerted, long-term effort to close them.

Gains Before and After NCLB

Many states had reforms well underway before NCLB, so it is useful to know whether the pace of improvement has picked up since NCLB took effect. Only 13 states supplied enough years of data to make this determination—too few to know whether the findings for this sample represent a true national trend. In nine of these states, test results improved at a greater average yearly rate after 2002 than before. In the other four states, the pre-NCLB rate of gain outstripped the post-NCLB rate.

Difficulty of Attributing Causes for Gains

This report focuses on whether test scores have gone up since the enactment of NCLB. We cannot say to what extent test scores have gone up because of NCLB. It is always difficult to tease out a cause-and-effect relationship between test score trends and any specific education policy or program. With all of the federal, state, and local reforms that have been implemented simultaneously since 2002, it becomes nearly impossible to sort out which policy or combination of policies is responsible for test score gains, and to what degree. In a similar vein, this report does not take a position on how well specific components of NCLB are working or whether the requirements in the current law are the most effective means to raise achievement and close test score gaps.

One more caveat should be emphasized: test scores are not the same thing as achievement. Although tests are often viewed as precise and objective, they are imperfect and incomplete measures of how much students have learned. Still, state tests are the primary measure of achievement used in NCLB and are the best available standardized measures of the curriculum taught in classrooms.

Need for More Transparency in Test Data

The No Child Left Behind Act requires states to report a massive amount of test data and attaches serious consequences to these data for districts, schools, and edu-

cators. But the data on which so much rests are not easy to access in some states and are sometimes inconsistent, outdated, or incomplete. Moreover, the data needed to calculate effect sizes or determine which subgroups were small or rapidly changing were unavailable in some states, even though these data are integral to all testing systems. Reasons for these shortcomings include overburdened state departments of education, ongoing corrections in test data, and technical or contractual issues with test contractors. These shortcomings are not necessarily the fault of state officials—who were generally cooperative in providing or verifying data when asked—but these problems complicated our efforts to reach definitive conclusions about student achievement.

It took many months of effort to gather all the data needed for this study and have state officials verify their accuracy. Our experience suggests how difficult it would be for the average citizen to get information about test score trends in some states, and points to the need for greater transparency in state test data. States could improve transparency by taking the following steps:

- Posting test data in an easy-to-find place on state Web sites
- Providing clear information and cautions about breaks in the comparability of test data caused by new tests or changes in testing systems
- Reporting standard deviations, mean scale scores, numbers of test-takers, and other important information listed in chapter 7 of our report.

State-By-State Achievement Trends on the Web

The trends highlighted in this testimony and in our report have been drawn from an extensive set of data on each state. Complete profiles of test results and other information for individual states can be accessed on the CEP Web site at *www.cepdc.org/pubs/stateassessment.* We encourage anyone who is interested in trends for a specific state to visit the Web site and find that state's profile.

Future Phases of This Study

This report describes the findings from phase I of what will be a three-phase study of student achievement. Phase II, which will be completed this summer, involves on-site interviews with state officials in 22 states. Phase II investigates in more detail the trends uncovered during phase I of the study and the factors that affect comparability or availability of test data; it also reports information from state officials about how well specific requirements of NCLB are working and how the law could be improved. Phase III, which will be carried out in the fall and winter of 2006-08, examines student achievement at the school district level in three states.

Thank you, Chairman Kildee, and we are available to answer any questions.

Chairman KILDEE. Thank you very much.

STATEMENT OF RICK MELMER, SECRETARY, SOUTH DAKOTA DEPARTMENT OF EDUCATION

Mr. MELMER. Chairman Kildee, Ranking Member Castle and members of the subcommittee, my name is Rick Melmer from the State of South Dakota. I have been serving as the chief school officer in South Dakota for 4 years. I am grateful for the opportunity to come and talk to you about a topic that I know you care deeply about and we certainly do in the State of South Dakota. I also represent the Council of Chief State School Officers and this topic has been high on our agenda, as you can imagine, over the past year and a half.

Since 2002, the No Child Left Behind Act has challenged schools, districts and States across this country to ensure the best education possible for students. Since that time I think you can all agree that States, schools and districts have made significant progress, both in implementing policies and, as Mr. Jennings' report suggests, improving achievement scores and closing achievement gaps. We are grateful for that news. We are excited to hear about it. And it is really proof that the foundation has been laid with the No Child Left Behind Act of 2002 for some good things

in the future. Our hope is that we can take advantage of this opportunity now to reauthorize this act and make it even better as we look ahead 5 to 7 years down the road.

The Council of Chief State School Officers has worked for the past 3 months in three primary areas: One, the development of a policy statement; two, eight specific recommendations in how the law can be improved; and, finally, specific legislative language that can be used when you are considering reauthorization. The purpose of my remarks today though is going to be focusing on three primary areas where flexibility could benefit States across this country.

The first area is the promotion of innovative models in reinventing the peer review process. We have all heard about growth models. In fact, there are seven States today that have been approved by the Department to implement growth models. There are many more States out there like South Dakota that would like to be one of those States. At times we are challenged, especially in rural States, to have the expertise we need to build the capacity to create growth models. It would be my hope that as we look ahead to the future that the NCLB reauthorization would include the ability to help States, all States if they want to reach a goal of growth models or other innovative testing systems that could benefit students in their State.

We spend a lot of time talking about growth models, but I hope we don't forget that there could be other ways to assess kids, especially when you are looking at 21st century skills, ways like portfolios, project-based learning and possibly even computer simulations that would test whether the young people of today have the skills they need to be successful.

The second part of my comment is reinventing the peer review process. Currently the peer review process is based almost exclusively on compliance. And frankly when the law was instituted in 2002 that was probably the right model to implement because the U.S. Department of Education had an obligation to ensure that every State complied and set the right foundation. The good news is almost every State in the country has been approved under the Standards and Assessment Program. So the need for the compliance peer review is less today than it was in 2002. So we are suggesting a peer review process that is a lot more transparent and a lot more collaborative in its efforts in working with the States.

You may know this, but we have no opportunity to even talk to our peer review members when they are reviewing our State's Standards and Assessment Program. The best peer review process is an interactive process where you can sit at the table and have dialogue about your concerns and issues and the peer review members can get a better feeling from a State as to what the issues really are.

The second area I wanted to emphasize is differentiated consequences. In South Dakota we have 700-plus schools in about 168 school districts. As you can imagine, some of our schools in our State do very well. We also have some that hit—for example, in our largest district of Sioux Falls, there are three schools in Sioux Falls that hit 17 of 18 of their academic targets under No Child Left Behind. And yet those schools are treated the same as a school that

hits 0 of 18 academic targets. We lack the ability at the State level of treating individual schools differently based on their academic performance. If you hit 18 of 18 you are treated the same as someone that hits 0 of 18. Differentiated consequences have to be a part of reauthorization as we move into the future.

And finally, teacher quality. We would all agree on the need for teacher quality and the importance of it. But in rural States like South Dakota where you have 45 districts with less than 200 kids high school teachers oftentimes teach many different disciplines and oftentimes five, six and even seven preps in a day. It is difficult to recruit high school teachers in our State when we are expecting them to take tests in three different content areas in five or six specific course content areas. So the Highly Qualified Teacher Act, even though I know its intentions are honorable and important, makes it increasingly difficult for us to recruit high school teachers and make sure that they are highly qualified in all other content areas.

I am grateful for the opportunity to testify today and look forward to the question and answer period that will follow.

[The statement of Mr. Melmer follows:]

Prepared Statement of Hon. Rick Melmer,
South Dakota Secretary of Education

Chairman Kildee, Ranking Member Castle, and Members of the Subcommittee, thank you for the opportunity to testify today on providing appropriate flexibility in the reauthorized Elementary and Secondary Education Act (ESEA). My name is Rick Melmer, and I am the Secretary of Education in South Dakota, a position I have held since 2003. I am also the president-elect of the Council of Chief State School Officers (CCSSO) and have previously been a local superintendent in Watertown, South Dakota and in Sioux City, Iowa.

Passage of the No Child Left Behind Act (NCLB) appropriately challenged states and school districts to redouble their efforts to ensure the success of all students. During the last 5 years, states have made tremendous strides in implementing the policies and programs needed to improve student achievement and close achievement gaps. In fact, every state has implemented state standards, state assessments, state accountability plans, and state teacher quality plans. Now the question is how do we use the opportunity presented by reauthorization to form a new state-federal partnership designed to build-on, and effectively use, the strong foundation laid by states and districts during the last few years. Working together I believe we can make a difference for each and every child in this country.

CCSSO and its members worked for nearly a year and a half in preparation for reauthorization by developing a policy statement, eight specific recommendations, and then legislative language that codified the eight recommendations into current statute. We also partnered with the National Governors Association and the National Association of State Boards of Education to deliver a separate, joint reauthorization statement to Congress.

CCSSO's membership believes that reauthorization should emphasize three principles (1) innovation, (2) capacity building; and (3) research and dissemination of best practices. These principles, and our specific recommendation for achieving them, have been delivered to the Committee through the testimony of nearly a half-dozen of my state colleagues who participated in recent hearings on adequate yearly progress, growth models, students with disabilities, and supplemental education services. I am here today to reinforce our recommendations and to talk about their specific importance in the rural context. I ask that you keep these concepts at the forefront of your internal discussions as you begin to craft the changes to the ESEA that will guide the education reform process for the next five years.

This hearing focuses appropriately on examining the proper level of state and local flexibility needed to ensure that the reauthorized ESEA helps, not hinders, the education reform process during the coming years. Congress must continue to hold states accountable for improving student achievement and closing achievement gap, while also providing them with the flexibility needed to implement innovative models for accomplishing these vital national goals. States are chomping on the bit to

move forward with creative, innovative solutions to many challenging problems, but the current framework is inhibitive and too rigid to recognize unique state and local challenges and opportunities, particularly in rural areas.

Given the pace of change and the dramatic improvements we need in student achievement to make every child a graduate ready for college, work, and citizenship, the question is: How do we build a federal law that promotes state action and innovation, with continuous improvement over time? Flexibility and support are core strategies to achieving this goal. But what we need is a culture shift in federal law. Flexibility should not be understood as bending the rules, but should rather be available whenever it makes the best educational sense for students. Innovation should be the hallmark of federal law, in which states are encouraged to build better education systems that improve student achievement in their particular state contexts and may provide promising models for other states. This approach is particularly important in small and rural states like South Dakota, but also to my state colleagues who are moving towards innovations such as the use of formative and web-based, embedded assessments.

How can federal law codify such innovation without undercutting core principles? We have several concrete recommendations that build on NCLB.

I. Promote Innovative Models and Reinvent Peer Review

States should be encouraged to implement innovative education reform models, so long as they can demonstrate, through a revised peer review process, that their approach is educationally sound and is designed to raise, not lower, the achievement bar. The new system must also better recognize when schools and districts are making real progress. Rural states like South Dakota know what needs to be done to move forward, and could benefit significantly by having greater flexibility to address the unique problems they face.

NCLB properly focused the nation's attention on improving basic skills for all students. Now the new law should take the next step forward by fostering a "culture of innovation." Implementing this new approach will require incentives for encouraging innovation and a transformed peer review process. The current peer review process is a challenge not only for states, but also the U.S. Department of Education. South Dakota is now in the middle of its standards and assessment peer review and is currently labeled as "approval pending." We admittedly have work to do to make our system better, but the current peer review framework does not always facilitate improvement. For example, a revamped peer review system supported by greater resources would enable the Department to provide more timely communication to us and to other states. The current process would also be more effective if it included a strong technical assistance component that informed our work. Rigid penalties are also a problem. For example, even though all interested parties acknowledge that we have made significant progress over the past year in improving the state's assessment system, 25% of our Title I administrative funds may be withheld. Such withholding will make it even more difficult for us to accomplish our objectives. I believe the new law should reward or acknowledge improvement, and avoid rigid penalties for states, like mine, that are making good faith efforts to improve.

In sum, we believe a revised peer review process should grant states a role in the selection of qualified peers, focus on technical assistance, full transparency, real communication and dialogue with states, consistency in peer review standards and outcomes across states, timeliness of feedback and results, dissemination of best practices, and more.

II. Improve Accountability Determinations

States should be able to use a variety of accountability models, including growth models and compensatory data that build on AYP, to promote more valid, reliable, educationally meaningful accountability determinations.

South Dakota applied to be a growth model pilot state for the testing year 2006. Unfortunately, the Department denied our application and many other states' applications. More states might have been able to take advantage of this important flexibility if the law placed a greater emphasis on fostering innovation and provided increased resources and strong technical support. For example, many rural states do not have the "in-house" expertise (i.e., psychometricians) to develop and evaluate their own assessment systems. In this instance, an enhanced peer review process that includes technical assistance and provides incentives for innovation could have had powerful results. Therefore, while we strongly urge you to encourage growth models as part of ESEA reauthorization, we also ask that you ensure the new law encourages states to pursue such innovations and provides proper financial and technical supports needed to help them succeed.

III. Differentiated Consequences

The reauthorized ESEA should encourage a full range of rewards and consequences for districts and schools that differ appropriately in nature and degree. Based, for example, on whether schools miss AYP by a little versus a lot. In that context, the new law should permit states to exercise appropriate judgment and differentiate both accountability determinations and consequences based on sound evidence.

This is particularly important in rural areas where the rigid consequences of NCLB often do not fit the needs of the school or district struggling to make improvements. For example, the Sioux Falls School District in South Dakota is our largest district. The Sioux Falls district is currently on Level 2 of District Improvement even though the district has reached over 80% of the academic targets (180 of 224). Furthermore, there are three schools in Sioux Falls that reached 17 of 18 academic targets and yet remain "in improvement." This designation is the same as a school that reaches 0 of 18 would receive. The "all or none" approach to school and district improvement must change to reflect an accurate assessment of educational progress or lack thereof. Unless a school in improvement reaches a perfect score two years in a row, the school remains "in improvement."

IV. Enhance Teacher Quality

Incentives should be put in place for states to create the best teaching force by continuously improving teacher quality, by supporting best-in-class professional development, and by using multiple individual pathways to pedagogical and subject matter expertise.

South Dakota has 45 school districts with less than 200 students in the K—12 districts. High school teachers are expected to teach in multiple disciplines in order for the small high schools to meet the state's graduation requirements. As a result, the highly qualified teacher guidelines, which tend to favor large districts with specialized teachers, can hamper a rural district's ability to meet the intent of the law.

The highly qualified expectations for high school special education teachers have made a challenging circumstance even more difficult. South Dakota currently has a shortage of special education teachers, especially at the high school level. The current law that requires a special education teacher to be highly qualified in all content areas is unrealistic and problematic in rural states.

I mentioned only a few of the areas where rural states like South Dakota have felt most challenged by the rigidity of the current framework. We have learned a lot in just the past few years about what is working in our schools and what is not. It's fair to say that the federal government, nor states, nor districts, nor schools have all the answers, so the law must provide room for continuous improvement and states and districts should be able to use their judgment about how to accomplish NCLB's core objectives.

Thank you again for the opportunity to testify today and I look forward to answering any questions you may have.

Chairman KILDEE. Thank you very much.

STATEMENT OF KATHLEEN STRAUS, PRESIDENT, MICHIGAN STATE BOARD OF EDUCATION

Ms. STRAUS. Thank you, Chairman Kildee, Congressman Castle and members of the subcommittee. I am delighted to be here and greatly appreciate the opportunity to appear before you and discuss flexibility in the most recent version of the Elementary and Secondary Education Act, or the No Child Left Behind Act. I am particularly pleased and honored to be here with Chairman Kildee in the position he is in because, as he said, we go back a long way to when he was in the State Senate in Michigan and I was a Staff Director for the Senate Education Committee at that time. So we are very proud to have you in your position and very proud of you.

And I am privileged to be here today to represent not only the State of Michigan as President of the State Board, which in Michigan is an elected statewide body, bipartisan, but I am also speaking on behalf of the National Association of State Boards of Edu-

cation and my colleagues who serve on State boards throughout the country.

I want to make it clear that Michigan State Board of Education and indeed all the State boards across the country embrace the philosophy and the goals of the No Child Left Behind Act. It is our belief that the fundamental aspects of the law are positive and well-intentioned.

As State education leaders we have championed the theory for many years that all children can learn. But it is also our belief of State boards generally, and in particular the entire State Board of Michigan, all eight members, Republicans and Democrats alike, that modifications to NCLB are necessary to reach these goals.

In the initial phases of implementation there was no aspect of the law that was more welcome than flexibility, nor more touted I might add. We were soon to learn, however, that the flexibility existed more in theory than in application. What we inherently knew as State board members at the State level and throughout the country was that we have 50 separate distinct State education systems. A one-size-fits-all approach is difficult if not impossible to apply throughout the country.

Speaking from personal experience, this became painfully clear as we in Michigan parsed through the law page by page all those many, many pages and provisions and tried to make it fit into the academic frameworks, assessment schedule and the accountability system that we had previously and so successfully established in Michigan. We came to the conclusion that while we are meeting the spirit of the law we clearly needed more flexibility to help our good faith efforts in meeting the letter of the law.

As a result, I am here today to reaffirm the NASBE recommendation, the National Association of State Boards of Education recommendation, that we need to move from a law of absolutes to one that incorporates the following principles:

One, provide adaptation in State assessment requirements, particularly for testing of special needs students, such as students with disabilities and limited English proficient students.

Two, to permit the use of growth model measures in all States.

Three, to provide accommodations in teacher qualifications, deferring to well-established State licensure procedures, recognizing in particular the challenges of staffing in rural areas and high needs subjects.

Fourth, to recognize the enhanced role of States in education leadership, technical assistance and school improvement with a solid, consistent Federal investment for State capacity that reflects the State-Federal partnership in improving low performing schools. All the States are providing a great deal of technical assistance but need the capacity to do so.

And fifth, promote fair, consistent and equal treatment in all dealings, negotiations and approvals between State and Federal officials supplemented by peer review, as Rick just said, consisting of accomplished, credentialed, well-trained professionals knowledgeable in State and Federal education policy and law.

As you know, these issues surrounding ESEA reauthorization are of such concern to State educational leaders that NASBE, the National Governors Association and the Council of Chief State School

Officers included these things, among others, in a recently proposed set of joint reauthorization recommendations that were submitted to Congress. Perhaps the most important suggestion I could make today on behalf of State policy makers is to give States that have served as the laboratories of innovation and reform the latitude to address their unique circumstances. States should be extended the freedom to development and implement policies that meet their specific needs while remaining within the spirit and letter of the law.

Admittedly, some areas have been addressed but clearly many more aspects need attention and collaboratively developed resolutions. In Michigan's accountability workbook submissions, for example, to the Department of Education that serve as our current day annual plans we have asked for such latitude. But I regret to say that a fair amount of what we have thoughtfully compiled and presented has been rejected, often after months of delay and sometimes having been accompanied in the first and subsequent instances by what we thought were encouraging commentaries of acceptance.

Unfortunately, our experience in Michigan has not been unique. As a State that is generally recognized as a national leader in education and as one of some 18 States that have received full approval for our assessment system, what would we specifically request.

You have our whole statement, but I will be glad to answer any questions you have. So thank you very much for this opportunity.

[The statement of Ms. Straus follows:]

Prepared Statement of Hon. Kathleen N. Straus, President, Michigan State Board of Education

Chairman Kildee, Congressman Castle and Members of the Subcommittee, please accept my sincere appreciation for the opportunity to testify today on flexibility in the most recent version of the Elementary and Secondary Education Act/No Child Left Behind Act. I am privileged to appear before you today, representing not only the State of Michigan as President of the statewide, elected, bipartisan State Board of Education, but also speaking on behalf of the National Association of State Boards of Education (NASBE) and my colleagues who serve on state boards of education throughout the United States.

Initially I want to make it abundantly clear that the Michigan State Board of Education and, indeed, all state boards of education embrace the philosophy and goals of the No Child Left Behind Act. It is our belief that the fundamental aspects of the law are positive and for the most part well intentioned. As state education leaders, we have championed the theory for many years that all children can learn. But it is also our belief—of state boards generally and the entire Michigan State Board of Education, all eight members, Republicans and Democrats alike in particular—that modifications are necessary to the amendments made in the 2001 reauthorization.

In the initial phases of implementation of the No Child Left Behind Act, there was no aspect of the new law more welcome than flexibility, nor more touted, I might add. We were soon to learn, however, that the flexibility existed more in theory than in application. What we inherently knew as state board members at the state level and throughout the country, was that we essentially have 50 separate, distinct state education systems. A one-size-fits-all approach is difficult if not impossible to universally apply throughout the country. Speaking from personal experience, this became painfully clear as we parsed through the law page by page and provision by provision, and tried to make it fit into the academic frameworks, assessment schedule, and accountability system we had previously and so successfully established in Michigan. We came to the conclusion that while we are meeting the spirit of the law we clearly needed more flexibility to help our good faith efforts in meeting the letter of the law. As a result, I am here today to reaffirm the NASBE

recommendation that we need to move from a law of absolutes to one that incorporates the following principles:

Provide adaptation in state assessment requirements, particularly for testing of special needs students such as students with disabilities and Limited English Proficient (LEP) students;

Permit the use of growth model measures in all states;

Provide accommodations in teacher qualifications, deferring to well-established state licensure procedures, recognizing in particular the challenges of staffing in rural areas and high-need subjects;

Recognize the enhanced role of states in education leadership, technical assistance, and school improvement with a solid, consistent federal investment for state capacity that reflects the new state-federal partnership in improving low-performing schools;

Promote fair, consistent and equal treatment in all dealings, negotiations, and approvals between state and federal officials, supplemented by peer review teams consisting of accomplished, credentialed, well-trained professionals, knowledgeable in state and federal education policy and law.

As you know, these issues surrounding ESEA reauthorization are of such concern to state educational leaders that NASBE, the National Governors' Association, and the Council of Chief State School Officers included these themes among others in a recently-proposed set of joint reauthorization recommendations submitted to the Congress.

Perhaps the most important suggestion I could make today on behalf of state policymakers is to give states that have served as the laboratories of innovation and reform the latitude to address their unique circumstances. States should be extended the freedom to develop and implement policies that meet their specific needs, while remaining within the spirit and letter of the law. Admittedly some areas have been addressed, but clearly many more aspects need attention and collaboratively-developed resolutions.

In Michigan's accountability workbook submissions to the U.S. Department of Education (USED) that serve as current day annual plans we have asked for such latitude. Some of what we have sought has been accepted. But I regret to say that a fair amount of what we have thoughtfully compiled and presented has been rejected, often however after months of delay, and sometimes having been accompanied in the first and subsequent instances by encouraging commentaries of acceptance. Unfortunately, our experience in Michigan has not been unique.

As a state that is generally recognized as a national leader in education, and as one of some 18 states that have received full approval for our assessment system, what would we specifically request? Let me briefly provide you with our priorities:

Graduation Cohorts of More Than Four Years Recognizing that time is the variable for some students to achieve the more rigorous graduation requirements recently adopted in Michigan and across the nation, we must have the flexibility to use graduation cohorts of more than four years under some circumstances. This is especially necessary for alternative education programs that accept and embrace students who are far behind grade level and are punished by the current system when they are unable to graduate the individual students with a four-year cohort.

Use of Best Score Through Grade 12 in Adequate Yearly Progress Calculations (AYP) Michigan would like to incorporate the student's best score, including senior retests, in AYP determinations. The best score for students in calculating high school AYP would be used through Grade 12. We recommend the use of alternate assessments measured against alternate/modified achievement standards based on individualized growth expectations across grade levels, as needed for some students.

Identification of School or School District for Improvement It would be preferable to identify a school or school district for improvement only if the school or school district does not make AYP for the same content area in the same subgroup for two consecutive years.

Proxy Calculation for Students with Mild to Moderate Cognitive Impairment Allow the "standard number of years" for graduation to be more than four under special circumstances.

Permit the Development of Appropriate Assessments for Students with Disabilities An assessment between the current "1 percent assessment" and the newly-permitted "2 percent assessment" would help states assure that all students with disabilities are assessed appropriately.

Limited English Proficient Students and AYP Allow schools and school districts to expand flexibility for English Language Learners (ELLs) in their first year of school in the United States to their first two years of school in the U.S. Allow ELL students to reach proficiency in English before testing in English; allow standard number of years for graduation to be more than four. Permit states to properly in-

clude new immigrant ELL students in school accountability, based on multiple measures for several years (no fewer than three), where educationally appropriate. Allow a full range of alternative assessments, and a system that values individualized growth. Recognize the positive performance of students who have recently transitioned out of the ELL student subgroup accountability determinations for an appropriate period.

Consistency with Approvals of Exceptions Among States In Washington, the current terminology is transparency. In Michigan we would refer to it as equity, fairness, and respect. In the creation of state plans and the approval of accountability workbook modifications, USED should maintain a policy of consistency. Uniformly sharing information about approvals openly among states would foster great mutual respect and trust, and at the same time assist states in resolving similar difficulties. Some examples of inconsistency have been approval of various N sizes, confidence intervals, and assessment of ELL students.

Thank you again for the opportunity to offer Michigan's State Board of Education perspective and that of our national association. I would be pleased to answer any questions you may have, or provide background information to support the issues I have raised today.

Chairman KILDEE. Thank you very much and your entire statement will be included in the record.

Miss Johnson.

STATEMENT OF CAROL JOHNSON, SUPERINTENDENT, MEMPHIS CITY SCHOOLS

Ms. JOHNSON. Good afternoon, Chairman Kildee, Congressman Castle, and members of the subcommittee. I am Carol Johnson, Superintendent of the Memphis City Schools. I have been in Memphis for 4 years and 6-1/2 years in Minneapolis, Minnesota before that as Superintendent. Thank you for the opportunity to testify on flexibility issues and on No Child Left Behind.

This school year I am really proud to say that the State of Tennessee declared Memphis City Schools to be in good standing under No Child Left Behind. It was the first time since No Child Left Behind has been put in place. In 2004, we had about 62 schools that were deemed a high priority by the State. And indeed we have a lot more work to do, but today we have about half that number, about 36 schools.

The Council of the Great City Schools in their Beating the Odds report reported that our school district was making faster progress than the State. And we believe that flexibility means different things to different people. And so my comments today will really focus a great deal on where we think the flexibility should occur.

But before I begin let me just say that I think that No Child Left Behind is at a critical juncture where failure to address concerns about the implementation threatens to undermine its original and noble purpose of creating academic success for all students. So I hope today, and as you review the bill, that you will certainly look at ways to enhance the flexibility that I think is desperately needed.

Today I would like to briefly summarize five key points: School intervention and improvement framework, a little bit about the growth model and data systems, transferability and staffing high priority schools.

We believe, and if you look on page 7 of my remarks you will see that the Council has outlined a modification of the school intervention and improvement framework. The chart attached to my testi-

mony on page 7 illustrates our proposal and how it compares with the current law. This revision will allow the accountability timeline for schools to begin immediately. It focuses on improvement and acknowledges a sustained change happens over a multi-year period, not one year at a time, and shifts the sanctions from cascading and changing every year to helping schools to stay focused on the improvement strategy long enough to see real results. And further in the red boxes, it separates, and I think this is consistent with Secretary Melmer's comments, it separates those schools that are pervasive and persistent failures to those schools that have maybe one subgroup having difficulty. It does retain the parental choice that will remain and it moots the effects of receiving late test data from the State and gives schools additional flexibility in the use of funds.

The Council's emphasis on good and best teaching strategies during this initial intervention period I believe is consistent with what we are seeing in Memphis that really works. The focus on restructuring strategies is something I am familiar with in my experience both in Minneapolis and in Memphis. Over the past three years we have restructured, or fresh started as we call it, eight schools in Memphis and we begin restructuring for additional schools this fall. Of the eight schools that have been restructured six now have made adequate yearly progress after failing to make AYP for 6 consecutive years.

The Memphis City Schools restructuring model is known as Fresh Start and before we decide to restructure a school we don't just rely on the test score data; we have a team of external examiners come in and work with us to look at all aspects of the data, including survey data and other things about the school climate. Our restructuring program begins first by replacing the principal and then we have flexibility in our collective bargaining agreement to hire teachers out of seniority order and as well to alter the compensation structure so that they can get rewarded for actual results.

The growth model, and of course Tennessee is one of those States that has had extensive experiences with the growth model, and we agree with most educators that a growth model should be incorporated into the accountability system. I believe that what teachers want is teachers want to get credit for showing progress with students who may come not being English speakers, but who teachers teach to read, write and think in English. That progress, though it may be significant, sometimes it is not enough for them to make the adequate yearly progress that is needed.

All of us are working to improve our tracking system so that we know what works and we know who is achieving. In Memphis we have a formative assessment system to monitor student progress and we use that every 6 weeks.

Transferability, we believe that we need greater flexibility, keeping Title I and Title III separate, since they are very student focused, but in the other categories asking for greater flexibility.

And then just finally we believe that it is really important that we are able to use a range of tools to incent and support putting and keeping the best teachers with our most vulnerable students. We are using Teach for America induction and mentoring pro-

grams, a teacher incentive grant to reward performance and connect good teaching practices with reward systems.

Thank you.

[The statement of Ms. Johnson follows:]

Prepared Statement of Dr. Carol Johnson, Superintendent of the Memphis City Schools

Good morning Chairman Kildee, Congressman Castle, and members of the Subcommittee. I am Carol Johnson, Superintendent of the Memphis City Schools. Thank you for the opportunity to testify on flexibility issues under the No Child Left Behind Act (NCLB), legislation that we have worked hard to implement.

The Memphis City Schools (MCS) is a large urban school district comprised of 191 schools and 118,000 students. Approximately 77 percent of our students receive a free or reduced-price lunch. We serve a predominantly African-American student population, but have a growing enrollment of English language learners who now number over 4,800 students from a variety of countries. Some 14.4 percent of our students are enrolled in special education programs, of which about 12 percent are gifted. We are very proud to serve this diverse group of young people.

We are also proud of the work of our administrators, teachers, and community leaders. They are striving every day to improve the academic achievement of our students. This school year, the State of Tennessee declared our district to be in "Good Standing" under No Child Left Behind for the first time. In 2004, we had 62 schools that were deemed "High Priority" by the state, i.e., in need of improvement. Today, we have about half that number—36.

Our academic gains, in fact, were highlighted recently in the Council of the Great City Schools' latest Beating the Odds report. The report not only recognized our progress but also pointed out that we are improving at a rate that far out paces statewide improvements. Nevertheless, we know that we still have considerable work to do.

I am pleased to be testifying today on the issue of flexibility under NCLB. Flexibility, of course, means different things to different people. To a school superintendent, flexibility can mean the ability to move human and financial resources around to meet specified needs. But it can also mean the freedom to give the wrong contract to an unqualified group. To a principal, flexibility can mean the ability to hire the team he or she wants in order to meet AYP targets. It can also mean the latitude to hire a workshop speaker he or she heard at a recent convention. To a teacher, flexibility can mean trying a new pedagogical technique. It can also mean closing the classroom door and doing whatever he or she feels like that day. To a state, flexibility can mean experimenting with alternative assessments for English language learners. It can also mean excluding those students by setting high N sizes. Or it can mean defining one's own definition of academic proficiency.

What gives flexibility its meaning and power is accountability, and the ability to hold people responsible for attaining expected goals—often in exchange for that latitude.

I am a strong believer in flexibility and the accountability that should accompany it. The Council of the Great City Schools on whose Executive Committee I sit also believes in this general principle. As a group, we continue to support NCLB and have developed a series of recommendations for its reauthorization that expands maximum flexibility while retaining strong accountability. We have also proposed ways to fix the law's operational problems, and shift funds into activities with greater promise for raising student achievement and narrowing achievement gaps. We have retained the overall framework of the Act, but have suggested modifying its internal operating gears so that its initial promise is better realized.

1. Proposed Intervention and Improvement Framework

I would like to take a few minutes to describe how the nation's urban schools would modify the "school intervention and improvement" provisions of the law. The chart attached to my testimony illustrates our proposal and how it compares with current law.

We propose that a school not making Adequate Yearly Progress (AYP) would begin school improvement planning immediately, rather than waiting another year. The school plan would have to focus on low-performing students, particularly those in low-performing subgroups. A school with large numbers of students who were not proficient would have a more extensive plan than a school with a lesser numbers or percentages of low-performing students. During this one-year planning phase, schools would have the flexibility to begin staff development immediately and the

latitude of using Title I funds to acquire necessary instructional materials or technical assistance.

We would then consolidate the current School Improvement I, School Improvement 2, and Corrective Action phases of the current law into a single, three-year school intervention and improvement period. This three-year period would allow a school to use its funds for well-researched instructional strategies that have been shown to raise student achievement—such as differentiated instruction, coaching, research-based reading programs, tiered interventions, benchmark testing, professional development, and the like. The school would be required to use up to 30 percent of its Title I funding for professional development, choice, and supplemental educational services, but would have the flexibility to fund these activities at a variety of levels as long as parents retain the option of transferring to another school or pick an external, private SES provider. We would follow this initial improvement period with serious but more differentiated consequences than the law currently provides.

This overall approach would have a number of advantages over current law. First, it would allow schools the time to pursue promising instructional programming under the direction of the school districts without changing activities each year in pursuit of the cascading sanctions the law now requires. Second, it would allow enough time for the instructional strategies to work before sanctions were levied. Third, it would give schools additional flexibility in the use of funds. Fourth, it would mute the effects of late test data from the states because the school's status would be determined for a multi-year period. Fifth, it would retain parental choice. Finally, it would keep the most serious sanctions but place them at the end of a process that was devoted to raising achievement and narrowing gaps.

The Council's emphasis on good instructional strategies during this initial intervention and improvement period is consistent with what we are doing in Memphis to raise student achievement, and what the organization has learned from its highly successful Strategic Support Teams. In Memphis, we use a series of strategies to assist and support our "High Priority" schools, including——

Districtwide Strategies for All Grade Levels

- Administrative leadership training
- School monitoring and "walkthroughs"
- Cross-functional instructional teams
- DATA (Directing Achievement through Accountability)
- Formative assessments
- Professional learning communities to sustain improvement and change
- Behavioral supports (Blue Ribbon Initiative)

Elementary School Strategies

- Literacy academy at selected schools
- Voyager interventions—Grades 2-5 districtwide

Middle School Strategies

- Read 180
- Striving Readers (eight schools)
- Increased honors—level courses
- Making Middle Grades Work (district implementation)

High Schools

- High Schools That Work
- Small Learning Communities (9th grade academies)

I also have made a number of organizational changes to increase support for students, teachers, and schools by establishing an Office of Academic Affairs, an Office of Student Engagement, an Office of Research, Evaluation and Assessment, and establishing a new associate superintendent's position to lead professional development.

2. Differentiated Consequences and Restructuring

The Council's proposals follow this initial period of intervention and improvement with a series of differentiated consequences, a concept that has received much attention as of late. We would distinguish between two types of schools: schools that persistently and pervasively fail to make progress with a majority of its students, and schools that fail to make progress with students who comprise fewer than half their students. Schools in the first category would be required, after a planning year, to comprehensively restructure or close. Schools in the second category would be required, after a planning year, to pursue a restructuring strategy that was more ex-

plicitly focused on the students or subgroups that were not making progress and staff members delivering services to them.

The first category of schools under our plan would warrant comprehensive restructuring or closure if they could not make any academic progress. The second category of schools would not necessarily warrant closure if the majority of its students or subgroups were making AYP targets or showing progress. These schools, instead, would have to focus their efforts and strategies—under the supervision of the district—on the students not making headway. We would cap the number of these schools in either category at a manageable 10 percent of all schools in a large district.

In the past three years, I have restructured eight schools in Memphis, and will begin restructuring four additional schools in 2007-2008. Of the eight schools that have been restructured, six have now made Adequate Yearly Progress after having failed to make AYP for six consecutive years.

Before deciding to restructure a school, our Memphis staff have to document the specific intervention and support strategies that have been implemented. If these measures prove unsuccessful, the district then contracts with an external group to conduct a management and instructional review of the school. The results of this review are used to determine whether restructuring is in the best interest of students. If restructuring is called for, then wed engage school staff, parents, and the school community to support the restructuring and reform efforts.

The Memphis City Schools restructuring model is known as "Fresh Start." Our program begins by replacing the principal of the identified school. The new principal is then given the authority to appoint a new administrative support team that will work together to interview and hire an entirely new faculty and school support staff. Teachers in "Fresh Start" schools are paid for two additional weeks of professional development—one before the school year starts and one later. Teachers in "Fresh Start" schools are eligible for financial bonuses based on the school's progress toward student achievement goals.

This overall restructuring approach is not easy or free of controversy, but it can be more effective. The Council's proposal also makes sense because it matches the sanction more closely to the severity of the problem without letting schools with small numbers of subgroups off the hook. Finally, this proposed approach more fairly balances an emphasis on instructional improvement and budgetary and programmatic flexibility with the need for strong accountability at the end of the day. That balance is out of kilter under current law because of its overemphasis on punishment and under-emphasis on what it takes to meet the Act's goals—good instruction.

3. Other Areas of Flexibility, Authority, and Reduced Restrictions

a) Growth Models. Virtually every commentator on NCLB suggests that the law include a growth model that would consider academic progress as part of the Act's accountability system. We agree with adding this feature to the law. Because not every state will want to use this flexibility, however, the Council recommends that school districts with the data capability be allowed to use an approved growth model from another state as part of that district's accountability system under NCLB. For example, Denver or Omaha could adopt the Tennessee or North Carolina model to assess progress and determine AYP.

We in Memphis have benefited from participation in the Tennessee Growth Model Pilot Program. It has given us a more accurate picture of the impact of the school's educational program on individual student academic growth. And it has given us better data to inform instruction. Still, Congress should know that growth models are not the panacea for long lists of "failing" schools if the models are based on a "universal proficiency trajectory" tagged to 2013-14. Less than a dozen schools made AYP using the Tennessee Growth Model. Even fewer did using the North Carolina Growth Model because both models are simply variations on the current status model and do not provide much credit for actual growth across the range of student achievement. The Council has made a number of recommendations for the "safe harbor" provisions that would give more credit for growth even if the school and students remained below the target proficiency levels. We think this would help improve flexibility.

b) Improved Data Systems. An essential component of any growth model is the state and local data system necessary to implement and support it. The Council suggests that local school districts have the flexibility to use up to 1 percent of their federal education funds for improving local data systems.

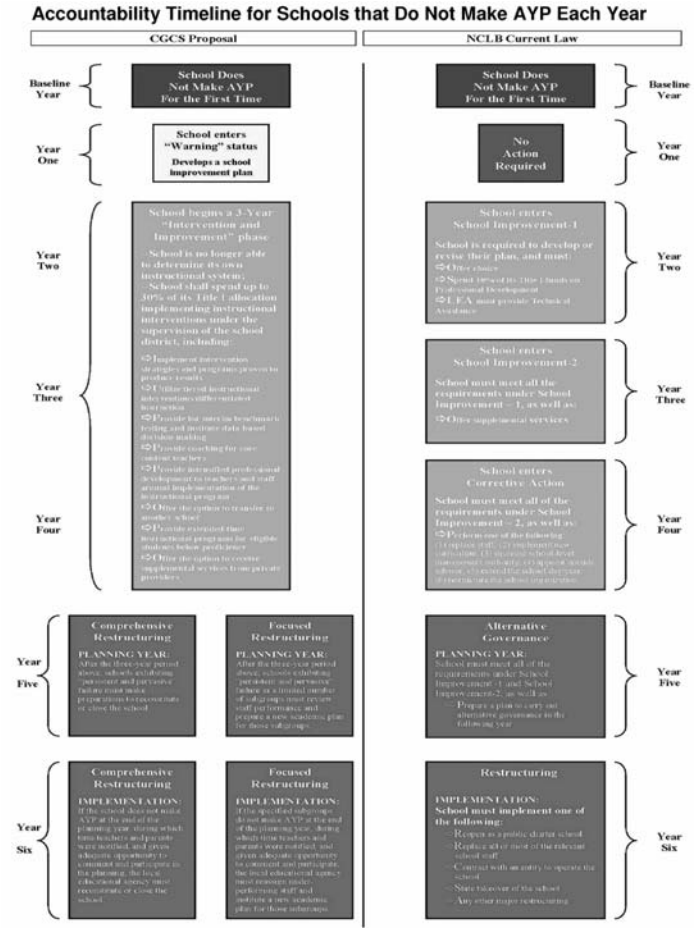
c) District Provision of SES or Extended Learning Programs. The Council has recommended retaining NCLB's SES program but proposes making it part of the schools' intervention and improvement program. With that change would come the

flexibility to use dollars on efforts that are more likely to boost the overall academic performance of children. Data collected by the Council also indicate that the numbers of participating students increase when the school district itself is a provider. The Council is urging that school districts be allowed explicitly the flexibility to provide those services.

d) Recruitment, Support, and Deployment of Staff in High Poverty Schools. The Council acknowledges that there is a serious national problem with the disproportionate placement of inexperienced teachers in high poverty schools. We would urge that school districts have the flexibility to use their ESEA funds for teacher recruitment, induction, mentoring, and other strategies to recruit, deploy, and support experienced and effective teachers in high poverty schools rather than mandating more requirements that schools cannot comply with.

e) Restrictions on ESEA Transferability. The Council proposed the transferability of ESEA funds in 2001 as part of the original NCLB authorization. We made this recommendation to allow school districts the flexibility to concentrate funding on a particular problem area while protecting the funding for the child-centered programs under Title I and Title III. Congress reduced this flexibility, however, when it limited the percentage of funds that could be transferred and further limited the flexibility for districts in improvement status under section 1116. Some school districts previously using funds for school improvement activities are currently being prohibited from continuing these initiatives. Moreover, regulatory restrictions from the Department of Education have discouraged districts from transferring funds into Title I. The reauthorization should remove the percentage restrictions and regulatory constraints and encourage the use of the transferability provisions for school improvement purposes.

I—like most of my urban colleagues—have supported No Child Left Behind from the outset, although I see all the same problems with the law that its detractors see. NCLB's focus on disadvantaged and minority student achievement is precisely the role that the federal government envisioned when it passed the original Elementary and Secondary Education Act in 1965. Unfortunately, it has taken until the beginning of the 21st century and the passage of NCLB for federal policy to get serious about the unconscionable achievement gaps that persist in our country. I hope that my comments today and the pragmatic recommendations from the Council of the Great City Schools will assist the Committee in revising the law in a way that will recapture the nationwide, bipartisan support that NCLB enjoyed at its enactment. Thank you.



Chairman KILDEE. Very good. We have votes on the floor of the House, but if you take 5 minutes we can make it over there and come right back, but we will hear from you first.

STATEMENT OF CHESTER E. FINN, JR., PRESIDENT, THOMAS B. FORDHAM FOUNDATION

Mr. FINN. Thank you, Mr. Chairman, Mr. Castle, members of the committee. It is nice to be invited. I feel as if I should be dragging a large national membership organization with me today to keep up with my colleagues, but in fact I am here on behalf of myself and my colleagues at the Thomas B. Fordham Foundation, three of whom came along today, including Mike Petrilli, who runs our national programs, recently spent 4 years at the Department trying to implement NCLB and currently is the Chief Editor of The Education Gadfly.

We believe that flexibility isn't a program or an amendment, it is actually a core principle, and it is a principle at the heart of the

most important question that Congress needs to answer with respect to the next go-around on NCLB. Namely, in K-12 education what should the Federal Government be tight about and what should it be loose about, when should Uncle Sam be proscriptive and when should he be flexible. I tend to view NCLB as a good first draft. And now you have a chance to edit it, revise it, hand it back to its students and make them improve what they were doing under the first draft.

How to fix it? Under the heading of the principle of flexibility these debates quickly become ideological. Conservatives tend to argue that States are in charge and the Federal Government should leave them alone. Liberals are apt to argue that States can't be trusted and that only strong Federal enforcement of specific actions will cause good things to happen. Neither of these views is right. Each leads to a bad outcome. The challenge is to strike an intelligent balance here.

I want to suggest that there are three guidelines worth following. First, whenever possible the Federal Government should be tight with respect to results and loose with respect to process and procedure.

Second, the Federal Government should figure out what it is actually good at and where it is apt to be most effective and only do those things, not try to do things it is not good at doing.

And thirdly, Washington should encourage States, districts and schools to earn more autonomy on the basis of strong performance and successful results.

My written statement elaborates on all three of those guidelines and offers a number of suggestions for specific measures that might be taken under each of them. If you all don't have to run, I will go over a few. If you want to run, I am happy to wait for questions.

Chairman KILDEE. You can finish up.

Mr. FINN. What does it mean to be tight as to results and loose as to process? This is of course Management 101. Any large, complex organization sets expectations for its units and then gives them freedom as to how to attain those expectations and what results to produce. And yet Federal policy has so often gotten it backwards, obsessing about process and actually paying minimal attention to results.

I think NCLB's architects in 2001 believed they had gotten this straightened out and that NCLB really was about results. But in fact it turned out to be backwards. NCLB turned out to be proscriptive with respect to a number of procedures and inputs in schools and actually surprisingly laid back about results leaving it to each State to decide what results it wanted and how to measure them. And we have seen a whole number of reports and studies, including JACS, but also NCES. And we see that State standards are incredibly variable, literally all over the place with respect to what States are expecting of young Americans. This is not good and it has led Washington instead to try to control interventions, teacher qualifications, a whole bunch of inputs and procedures. And that part is not going very well. What would work far better is for you all to be quite proscriptive with respect to standards and tests and use that as an opportunity then to rein in the regulatory and super-

visory impulses of the Federal Government with respect to how schools ought be run and staffed and intervened in and operated.

I think, for example, spending restrictions should be lifted in return for results. I think that Mr. McKeon's thought on this point is spot on. I think staffing restriction should be lifted with respect to highly qualified teachers. It is far better to focus on student achievement than to focus on teacher credentials.

And third, the NCLB is very proscriptive in terms of the interventions that districts are supposed to make in schools year by year by year. It would be far better to let those who actually have to engage in the interventions figure out what sequence and what timetable is likely to work best as long as they are all being held to account for performance against a common timetable and results which in turn are illuminated by an enormous amount of sunlight, comparable sunlight, that everybody's results with respect to each other can be seen and observed.

I have got a number of examples under my second maxim, figure out what the government is good at and only do those things. And I have got a bunch more examples under the third maxim that you should encourage States, districts and schools to earn greater autonomy on the basis of their performance.

This might be the most novel point we are making here today. A number of cities have figured out that successful schools ought to actually have greater freedom to continue to succeed. This is a principle that could be applied to States and districts as well. The better they do, the more freedom they get to innovate and do things their own way and the less they have to conform to process and input requirements or regulations.

I could, as you know, go on but in deference to your clock as well as your vote, I would be happy to answer questions. Thank you very much.

[The statement of Mr. Finn follows:]

Prepared Statement of Chester E. Finn, Jr., President, Thomas B. Fordham Institute and Chairman, Koret Task Force on K-12 Education, Hoover Institution, Stanford University

Chairman Kildee, Congressman Castle, members of the subcommittee: thank you for the opportunity to testify today. I am heartened to see you take up the issue of "Current and Prospective Flexibility Under No Child Left Behind," though I am also aware that this is one of the last hearings currently scheduled before you start work on legislation. I hope it's a case of saving the best for last, not some sort of afterthought.

That's because "flexibility," properly conceived, shouldn't be considered an "add-on," a separate program, or a sideshow. Rather, it is at the heart of the most important question the Congress must answer with respect to the next iteration of NCLB and ESEA. Namely: in elementary-secondary education, what should the federal government be "tight" about, and what should it be "loose" about. When should Uncle Sam be prescriptive, and when should he be flexible?

A few weeks ago, at a National Press Club panel, Hartford Superintendent Steve Adamowski commented that "high-achieving organizations eventually, in some way, get this right: what do you hold tight, what do you hold loose." As a local superintendent, he wrestles with this question all the time. What should be done uniformly, with strong central office control? And what should be delegated to principals? Too much flexibility and schools can founder. Too much prescription and innovation is sunk.

In Washington, these debates about prescription versus flexibility and the proper federal role quickly become ideological. Conservatives tend to argue that states have constitutional authority for schooling and the federal government should simply leave them alone. (Never mind that plenty of states have an abysmal record of pro-

viding a decent education, especially for poor and minority kids.) Liberals are apt to insist that states can't be trusted and that only strong federal enforcement of specific measures will lead to a narrowing of the achievement gap. (Never mind that plenty of states were making decent strides in raising achievement and narrowing gaps sans federal prodding.)

Neither view is right. Each leads to a bad outcome: Either "put the money on the stump," let states and schools do whatever they want, and hope for the best; or micromanage fifty states, 15,000 districts and tens of thousands of schools through miles of red tape. Neither approach works, not, at least, if stronger student achievement is the metric by which success is judged.

Is there a way to transcend these tired and predictable arguments? Let me propose three pragmatic rules to determine when Uncle Sam should be "tight" (i.e., prescriptive) and when he should be "loose" (i.e., flexible):

1. Whenever possible, the federal government should be tight about results and loose about process.
2. The federal government should figure out what it's good at, where it's most apt to be effective, and only do those things.
3. The federal government should encourage states, districts, and schools to "earn" even more autonomy on the basis of strong performance.

Let's take a closer look and consider what these rules would mean for a revamped NCLB.

Rule #1: "Tight" as to Results, "Loose" as to Process

This principle comes straight from Management 101: excellent bosses give their employees clear direction and specify the results to be achieved. But then they cut their charges plenty of slack to use their own creativity, innovation, and resourcefulness to achieve those results as they see fit. In a corporate setting, CEOs are "tight" about the bottom line, but "loose" as to how a particular unit achieves it.

This idea is the driving force behind the past twenty years of standards-based reform. It's related to former Vice President Gore's efforts to "reinvent" government. It's standard practice in large organizations around the globe. It's also the essential theory behind site-managed schools and charter schools.

And yet in federal education policy, we usually get it backwards. We obsess about process and pay minimal attention to results.

NCLB's architects believed, I think sincerely believed, that they were straightening this out, that NCLB was, above all, about results, with plenty of interventions and sanctions for those states, districts, and schools that didn't produce them. But they didn't get it right. I would even say they made a 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 in the ways that each judges best, they instructed states to define "proficiency" in reading and math as they saw fit—and then got very prescriptive about timelines, calculations of progress, and year-by-year interventions.

Instead of regulating ends, in other words, Washington once again found itself regulating means, prescribing a hundred aspects of what states and districts should do when, by their own lights, their schools don't do an adequate job. That means way too much regulation on the one hand and, on the other, plenty of incentive for 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 enterprise of standards-based-reform in peril.

The surest way to end this such questionable practices—and keep Washington from playing a cat-and-mouse game with recalcitrant states—is to move to a system of national standards and tests, while simultaneously freeing states, districts, and schools to achieve those standards as they see fit.

To be very clear, federal officials do not themselves need to, and in my view should not, create such national standards and tests themselves. But the federal government could require or encourage their use.

What does this have to do with flexibility? Perhaps counter-intuitively, I see national standards and tests as an opportunity to rein in Uncle Sam's more dictatorial and bureaucratic impulses. For forty years, Washington has sought to improve schools by regulating what they do. NCLB's mandated cascade of interventions into low-performing schools, for example—a different one each year for seven consecutive years—illustrates this pattern of behavaior. (As for the parallel cascade of state interventions into low-performing districts, the less said the better. It's a complex

mandate that may best be described today as "ignored".) Another example: by requiring testing in just two subjects and resting its entire intervention-and-accountability edifice on those results, No Child Left Behind has exerted definite pressure on schools to restructure their curricula, emphasizing math and reading skills to the detriment of other subjects. To date, we can find scant evidence that this strategy works—and some hints that it's backfiring. Schools are narrowing their curricula, neglecting already-proficient kids, lowering standards, and finagling test results. Common standards and tests would allow Uncle Sam to back away from his top-down, regulatory approach and settle instead for clarifying the objectives to be achieved, then measuring (and publicizing) whether states, schools, and students are in fact meeting them.

If states were in fact willing to sign up for tougher national standards and tests, what process-type regulations would I be willing to trade? Here are three categories:I21• Spending restrictions. School principals rightfully want control over their budgets. Yet current federal policy sends dollars into a myriad of silos, categorical programs that may or may not meet the needs of individual communities. NCLB's "transferability" provision began to address this problem by allowing states or school districts to shift funds from one silo to another, or into Title I. But it set a cap at 50 percent. President Bush and Congressman McKeon have it right when they call for expanding transferability to 100 percent, allowing states or districts to send all of their dollars into the Title I program and then ignore all rules and regulations for the other programs. This will cut red tape while also driving more federal dollars toward the needy students who need them the most. (I also favor Mr. McKeon's call to expand eligibility for "schoolwide" programs within Title I.)I21• Staffing restrictions. The impulse behind No Child Left Behind's "highly qualified teachers" provision is understandable. Teacher quality matters a lot, and most states have set miserably low standards for incoming teachers. Still, the mandate has created oodles of unintended consequences that need addressing. Fundamentally, it's worth asking whether the federal government should concern itself with teacher credentials or should stay focused laser-like on student learning. I prefer the latter. (For a compromise idea, see below.)I21• The "School Improvement" Timeline. Sure, states should take action when low-performing schools fail to improve year after year. What's not clear is whether NCLB's rigid sequence of prescribed annual interventions (including choice and tutoring, corrective action and restructuring) is any better than those that states might devise. In my view, such actions are far likelier to succeed if decided as close as possible to the problem and on timetables that make sense to those who will be responsible for implementing them. Moreover, ample sunlight shining down on school/district/state performance vis-a-vis clearly specified national standards will give state and local officials (and voters, taxpayers, parents, etc) good information by which to repair their own schools.

Rule #2: The Federal Government Should Figure Out What It's Good At, Where It's Most Spt to be Effective, and Only Do Those Things

Another key pragmatic question is whether Washington itself has the capacity, the infrastructure, and the know-how to implement NCLB's lofty expectations and detailed plans in an effective manner. Regrettably, the evidence is overwhelming that it does not. Nor will a change in Administration make much difference. That's because of a structural flaw in U.S. education federalism that NCLB inherited from earlier rounds of ESEA.

Back in 1965, when lawmakers' main goal was to disburse federal dollars to schools for additional instructional services for poor kids, it made sense, indeed was practically inevitable, to hand those dollars down the familiar institutional ladder from Washington to state education agencies to local education agencies. That was how state and local monies already flowed and there was no reason to create another mechanism to move federal funds. While SEAs and LEAs weren't always diligent in following Uncle Sam's rules, it was in their interest to comply, if only because they and their schools then got the money, which came without so many strings as to disrupt what they were already doing.

Today, however, getting Washington's dollars to the right places is the lesser mission of ESEA/NCLB. The law now deploys its funds and their attendant conditions, regulations, state plans, and oversight mechanisms to transform the system in fundamental ways, above all to boost student achievement and hold schools (and districts and states) to account for whether or not they accomplish this.

Thus arises a great paradox: Washington still relies primarily on SEAs and LEAs to do its bidding, yet now the point of federal programs is not to "help" them do more but to change what they do, often in ways they don't much want to be changed. In ways they judge contrary to their own interests. Ways that include ad-

mitting failure. And ways they may not be competent to handle, albeit ways that the public interest demands.

Why do federal policy makers assume that the very agencies that caused the system's problems (or, at least, allowed them to fester) now possess the will and capacity to solve them? The truth is, Congress and the White House never gave this any thought. At least I don't think you did. When crafting NCLB, I believe the craftspeople simply clung to the assumption that has ruled ESEA for four decades: that working down the familiar food chain is how Washington does business in the K-12 sector.

Thus NCLB proceeds in the accustomed sequence, with Uncle Sam telling states what to do, states telling districts, and districts doing most of the work. That hierarchy remains the basic architecture of federal education policy today as in LBJ's time. But its engineers never pictured it supporting a results-based accountability system, making repairs to faltering schools, or functioning in an education environment peppered with such disruptive, non-hierarchical creations as charter schooling, home schooling, and distance learning. It's as if a high-tech firm was officed in an old foundry without anyone bothering to re-wire, re-plumb, or even fumigate the structure.

This problem begins in Washington. Let's consider what NCLB has taught us about federal capacity:

1. The federal government is not good at nuance. Consider the law's complicated accountability and AYP provisions, for instance. The various design problems are legion, but they exist because of the principle that states must all be treated the same. Because some states were considered to be untrustworthy and unwilling to hold their schools accountable, especially for the performance of poor and minority students, all states were treated with suspicion. Thus the decision to mandate required elements of AYP, rather than setting broad parameters, which has led to constant cries for more flexibility. When laggard states complain about these prescriptive requirements, it's easy to label it "whining." But when leading states with well-developed accountability systems complain too, it's a sign that the federal hammer might be breaking some things that weren't previously broken.

2. The federal government can force recalcitrant states and districts to do some things they don't want to do, but it can't force them to do those things well. Yes, Uncle Sam has had plenty of practice at the compliance game and, on issues that are black or white (are states testing all students as required, for example), it can intervene and even take away dollars from misbehaving jurisdictions. But most of the important parts of NCLB are gray zones. Take "highly qualified teachers" or "public school choice" or "restructuring." In each of these areas, we've seen states and districts go through the motions without actually living up to the spirit of the law. Yet Washington is toothless to do much about it. That's not a legislative failure, it's a fact of organizational life. The federal government doesn't run the schools or employ their teachers; it has limited ability to make these complicated functions go well. But "going through the motions" isn't enough if we actually want to transform schools, and it fosters more cynicism.

3. States and districts do respond to carrots. What the federal government is actually good at—beyond distributing money, collecting statistics, investigating specific wrong-doing, and doing research—is funding promising reforms via competitive grant programs. Consider the Teacher Incentive Fund, for example. While controversial in some eyes, it has spurred several large school districts to experiment with merit pay for teachers. Something that would not have happened, in all likelihood, without federal dollars. Or look at the decade-old federal Charter School Program, whose funds are targeted to states with decent charter school laws. There's little doubt that federal leadership (first from President Clinton) played a key role in the charter movement's development. (That the charter program needs a makeover doesn't detract from the difference it has already made.)

What lessons should Congress take from NCLB's experience with federal capacity? First, even if you choose to continue to prescribe specific policies (such as AYP or Highly Qualified Teachers), aim for being clear about the ends and loose about the means. Take accountability, for example. If you don't accept the virtues of national standards, at least be more flexible about states' accountability systems. Rather than prescribing the exact nature of AYP, offer key design principles instead.

Let states prove that their systems measure up. Secretary Spellings' growth model pilot is a good example here. While she published a clear set of design principles and made states engage in a rigorous screening process, she didn't mandate a single uniform approach to measuring growth. Not doing so makes a lot of sense.

The second big lesson is that, if you want to see movement in a particular area, consider offering dollars to willing states and districts rather than mandating a

course of action for the entire country. When it comes to school choice, for example, adopt a version of President Bush's recommendation for a grant program for cities interested in expanding choice options, rather than forcing all 16,000 districts to go through the motions of offering choice when it's perfectly obvious that many of them lack the capacity as well as the will. Or when it comes to "Highly Qualified and Effective Teachers," look to Education Trust's recommendation to offer willing states extra cash to experiment with a "value added" system for measuring teacher quality, rather than adopting the No Child Left Behind Commission's suggestion of a nationwide mandate.

Rule #3: The Federal Government Should Encourage States, Districts, and Schools to "Earn" Even More Autonomy on the Basis of Strong Performance

"Earned autonomy" is an idea whose time has come. Increasingly superintendents (in Chicago, Las Vegas, New York City, etc.) are allowing schools to apply for greater freedom from central office. Those with a track record of improving student achievement qualify.

This same idea has made inroads in the charter-school domain. While charters have always been about "accountability in return for autonomy," increasingly their sponsors (including my own Fordham Foundation) understand that autonomy is something to be granted carefully. Once upon a time, some of us in the charter movement thought we should plant as many seeds as possible as quickly as possible and let a thousand flowers bloom; after all, we could always close them down. It turns out that closing schools is far harder than we thought. And we've witnessed many charter schools founder (or worse) because their leaders weren't prepared to work with the autonomy they had been given. So now conscientious sponsors screen applicants very carefully, just as venture capital firms screen prospective business start-ups. And only when a founding team proves that it is worthy of a charter and the concomitant autonomy is the green light given. We also reward charter schools for good performance by granting longer charters, hassling them less, and encouraging their replication. (The leash is shorter for low-performing schools.)

The appeal of this idea is obvious: It encourages good behavior (especially improved achievement), it recognizes that some entities are more capable of handling autonomy better than others, and it minimizes risk.

How could this principle be imported into federal policy? Here's what it might look like:I21• Grant greater flexibility to states that sign up for rigorous national standards and tests, or put their own rigorous system in place. As explained above, this flexibility could include expanding the funding "transferability" provision, waiving the school improvement timeline, etc.I21• Allow states with vigorous interventions greater AYP flexibility. Rather than trying to prescribe the exact sort of overhaul that states should serve up for failing schools, reward states that are engaging in effective reforms by giving them more leeway in defining their accountability metrics as they see fit. For instance, states that energetically provide school choice options to kids stuck in failing schools—by creating new charter schools for them to attend, or mandating inter-district transfers, or in other ways—might be allowed more discretion to differentiate sanctions for truly abysmal schools versus merely mediocre ones.I21• Allow schools that make AYP to ignore HQT. This is a particularly powerful idea. Improving teacher quality is necessary condition for boosting student achievement. But even the supporters of the "highly qualified teachers" provision admit that it's a poor proxy for school quality and classroom effectiveness, and that it's overly focused on paper credentials. So reward schools for getting great results by allowing them greater flexibility around staffing. To continue making AYP, schools will continue to make good decisions around teachers, but with less red tape from Washington. (This is especially important for high-performing charter schools, which are supposed to be freed from regulations in return for results, but are wrapped in the law's subject matter and certification requirements just like everyone else.)

Conclusion

You may have entered this hearing room contemplating some kind of new "flexibility program" for NCLB. I'm here to urge you to think more broadly, to ponder just where the federal government should be prescriptive and where it should be flexible. I hope you consider some of my specific proposals. I believe four of these have particular merit:

1. Encouraging states to adopt rigorous national standards and tests, and in return granting them greater flexibility around spending (by expanding "transferability ") and staffing (by waiving "highly qualified teachers ").

2. Moving federal requirements for state accountability plans away from prescriptive and pre-determined actions to more open-ended design principles. Be clear

about the end-result you want, accountability-wise, but flexible in terms of the specifics.

3. Reducing the number of mandates on states, districts, and schools, and instead offering competitive grants to entities willing to experiment with promising practices. In this category I would even include the law's "public school choice" provisions.

4. Allowing schools that make AYP to ignore HQT. At the end of the day you care about results, and good schools will ensure high-quality teachers. This show of goodwill and flexibility might go a long way.

Thank you once again for the opportunity to testify. I look forward to your questions.

Chairman KILDEE. Thank you very much. We will be right back. We have two votes. We will be right back.

[Recess.]

Chairman KILDEE. We should have an hour at least now without interruptions. Thank you, Checker, for finishing your testimony.

Jack, you discussed the need to improve the quality of data under NCLB. Can you discuss what data a State would need to collect, to implement and evaluate a sound growth model?

Mr. JENNINGS. I have to make a correction in my statement. It is page 82, not page 81.

Chairman KILDEE. That is all right.

Mr. JENNINGS. Alice Cain, Congressman Miller's staff person, said what page can we rip out.

Chairman KILDEE. That will be on the record now.

Mr. JENNINGS. We have issued a report several weeks ago about State departments of education. And we think of State departments of education as the agencies that carry out Federal law. But they are primarily State agencies and they have enormous responsibilities that we don't pay much attention to from the Federal level. And these State departments of education are severely restrained in terms of person power, in terms of funding. They just need much better support if they are going to help improve education. In fact, we find local school districts go to State departments of education more than any other agencies and yet they are handicapped because they don't have enough personnel. And what is happening is State departments of education are being converted into assistive agencies where they are starting to help local school districts more to bring about improvement and they need help with this transition. So we recommended in this other report that States get just an encouragement grant from the Federal Government because they are State agencies and have State leaders to help rethink State departments of education because they are finding all sorts of problems like with data. What is happening is if they get some good data people the technology companies come in and hire those people away because they can pay them more money. And yet from the Federal level and even at the State level legislatures are telling State departments of education to have better data systems, collect more data, make more data available to the public, but they are severely restrained in trying to do that.

So I would urge the Congress to pay some attention to the condition of State departments of education as the agencies that provide all this data and help them get up to the task, and that includes the growth model. The growth model if it is enacted, and it sounds like there is considerable support for it, is going to require the gen-

eration and use of a large amount of data. And if it is going to be used intelligently the States need help in using that data, but also local school districts and teachers are going to need help in interpreting that data to use it to the best effect.

Chairman KILDEE. Right now under AYP we test at, say, the third grade in school A and then the next year we test the third grade but they are different students because the third graders are now for the most part fourth graders, and we say that the third grade has not reached AYP. To take one form of the growth model you would actually follow the child and see how much that individual child has grown. Is that feasible or possible to have a growth model where you actually see how much each individual has grown to determine whether the school is making progress?

Mr. JENNINGS. Yes, that is feasible. However, it costs money. Because you have to have data systems, computer systems, and so on, you have to have identifiers for students, you have to be able to follow the students through their career. But if we really think education is important we should pay attention to every individual student and try to help every individual student. And the best way to do that is to be able to follow that student as the student goes through his or her own career. And it is a much fairer way to judge a teacher's performance to see how they have done with individual students as they go through school.

Chairman KILDEE. Health care seems to be ahead of education on that, is it not? While not perfect, you can generally follow the patient and get the records and follow better than what we do with students?

Mr. JENNINGS. Well, the trend now in health care is for doctors to have hand-held computers and bring up the records of their patients as they are visiting them and be able to go through all their records as they visit them. Teachers should be able to do the same thing. They should be able to use hand-held computers, use other technological advances and help kids in individualized instruction.

One of the problems with all this accountability is that we are generalizing everything with accountability tests that we are not paying attention to the individual children to help them improve, and we have to rethink that.

Chairman KILDEE. How much time do I have? I will ask one more question here. We have the NAPE test and each State has their own standards and their own test. And while probably this Congress would never want to apply the NAPE test across the board to every State to every student, can we use the NAPE test to test the State test.

Checker, do you want to tackle that?

Mr. FINN. Yes, sir. The new NCES report out today does a version of that using NAPE to compare State cut-off scores, State proficiency levels, in fourth and eighth grade reading and math. And it gives us a clearer calibration than I have ever seen before of relative levels of difficulty in State expectations on their own tests compared to NAPE. I think this kind of thing should be done all the time. If we are not going to have a national standard, which I think would be preferable, then at the very least we ought to have a whole lot of visibility of just how hard is Ohio's fourth grade standard versus Indiana's versus NAPE's.

Chairman KILDEE. Thank you very much. Anyone else have any comment on that? I defer to the gentleman from Iowa, Mr. Loebsack.

Mr. LOEBSACK. Thank you, Mr. Chairman. Just a couple of questions actually, Mr. Jennings. Not so much a question. I guess a comment. I appreciate the fact that you said we should be very careful about some kind of causal relationship. I am a former college teacher. I used statistics in my dissertation. That doesn't make me an expert by any means, but I am very aware of not drawing some kind of causal relationship. Just because some things are associated or whatever the case may be and using only 13 States out of 50 States, I am just very cautious about drawing conclusions. So I appreciate your caution as well. And that is just a comment more than anything else, but would you like to elaborate a little bit?

Mr. JENNINGS. Well, let me point out that we drew conclusions about pre and post-NCLB results from 13 States because they were the only ones that had data. However, we had data on 50 States and we had varying amounts of data on 50 States. On 50 States we had proficiency data. On 41 States we would have proficiency data and elementary. On 48 States proficiency data and something else. So we had large amounts of data on many States. But what we did was very strictly apply rules so that we had comparable data across years. So we would eliminate some States if they changed their tests. Thirty-seven States since 2002 have changed their tests in some way or another, either adopted a new test, put in a different cut-off score or whatever. So we are very careful to use comparable data across States.

Since you have somewhat of a scientific background or a research background, whatever you do you get criticized.

Mr. LOEBSACK. Of course.

Mr. JENNINGS. On this we tried to be purer than pure and make sure that everything we said was sound, and then we were criticized because we didn't use data that would have made us less pure because the data wasn't comparable.

Mr. LOEBSACK. I think you stepped into a huge minefield just by trying to determine whether NCLB has had any effect or not.

Mr. JENNINGS. One of the reasons we did that is that that is the important question. And we felt that if we didn't try, and this was an ideologically mixed expert panel that had varying points of view, that had deep expertise, we got nearly $1 million that we used for this, we did it for 18 months, we got the cooperation of 50 States, we felt if we didn't do this and try to answer that question then anybody in the world could stand up and give their opinion without any necessary data and say whatever they wanted to say.

Mr. LOEBSACK. And they still will, as you know.

Mr. JENNINGS. They still will, but I have been around for a long time and I felt that it was our duty to try to answer the question in the best way possible and put it out there. I hope it is not misused. But we tried our darnedest.

Mr. LOEBSACK. I appreciate it. Thank you.

Secretary Melmer, as was already said, I am from Iowa, not much bigger than South Dakota, but we only have 3 million people and my wife was a long-time second grade school teacher. And the

issue of teacher quality obviously came up in our house quite a bit and with a lot of her friends. And I could hardly go to a social event without hearing about NCLB and all the rest. That is why I am on this committee, by the way, in large part.

But the whole teacher quality issue, can you elaborate on that, because obviously places like South Dakota, States in particular that might have smaller populations, a smaller population base, maybe rural areas, it is very difficult, is it not, to attract quality teachers? And if that is the case then how—I mean the challenges it seems to me presented by NCLB are just that much greater. Is that true?

Mr. MELMER. Yes, it is true. As I mentioned in my testimony, 45 districts with less than 200 students in the entire school district, which means that you are talking about high school teachers that have to teach more than one discipline, in some cases three disciplines, depending on their background and preparation, and then you turn that into probably three to four to five different what we call preps—your wife would be familiar with that term—and that just makes it very, very difficult to be highly qualified in all of those areas.

At the same time we are having a hard time recruiting high school teachers anyway to come to a rural area to be paid a salary that sometimes some people would say is substandard, and then to throw on a bunch of additional requirements is challenging. At the same time we don't want to run away from the idea. We want our teachers to be prepared and ready to go. There just may have to be some consideration given to a waiver or some sort of a provision for a State that is really struggling to make all ends meet to allow us to continue to do the best job we can without necessarily having to follow all the letters of the law.

And Mr. Finn referenced the idea of if our results are good, maybe that should be some dispensation to say you don't have to follow all the guidelines if you are getting all the results.

Mr. LOEBSACK. My time has expired, but I just want to make one last comment. I was happy to see that the Iowa Department of Education did get its growth model approved by the Department of Education recently. But thanks to all of you for being here today. I appreciate it.

Chairman KILDEE. The gentlelady from Illinois, Mrs. Biggert.

Mrs. BIGGERT. Thank you, Mr. Chairman. Dr. Johnson, can you tell us what the percentage of your administrative paperwork is attributable to the Federal education requirements and what to the State and local education requirements? And then what portion of the Federal paperwork that your staff has completed is related to student performance? Everybody is always complaining about all the increased paperwork for No Child Left Behind.

Ms. JOHNSON. Congressman, members of the committee, I did not come prepared to calibrate for you exactly how much paperwork is involved in both either the Federal or the State. I will say that we certainly do a great deal of paperwork associated with submitting reports based on the report we have to give to the State for No Child Left Behind. And then because we have schools that have been a high priority we have to submit plans for each of those schools. Now, of course we would be doing that kind of account-

ability reporting about each school's improvement plan probably with or without No Child Left Behind. But I think that there are some provisions that we have to report on, not just for regular ed, but I think especially special education where our staff would say the paperwork consumes a great deal of the time.

Mrs. BIGGERT. Secretary Melmer, could you address that?

Mr. MELMER. In terms of the amount of paperwork, well, I have been in the State education agency for 4 years, and I am being told that it is by far more today than it was prior to my arrival at the department. In terms of the amount of time that we spend, the volume, I am unsure about that as well. I always have to be a little cautious about estimating because I am afraid I would be wrong. But I also understand because we at the State level do it to our local districts, I would assume if you had one of the superintendents from South Dakota here he or she would say the State gives us way too much paperwork and we don't have time to do all of it.

I understand at the Federal level if you expect results then you need to expect accountability to go with that. And as long as the Federal Government is continuing to provide dollars for State education, agencies are going to expect some accountability in return. So we try to balance all that out. But I think it is safe to say that the amount of Federal and State paperwork has increased over the last 3 or 4 years.

Mrs. BIGGERT. Dr. Johnson.

Ms. JOHNSON. I want to add one comment because if you looked on page 7 where we had revamped the school improvement, one thing that I think is a savings in paperwork, the way it is done now every year you are doing another school improvement plan as if you have got to start all over. The 3-year planning actually assumes that you are working with the same plan trying to make it better and you are focusing on those goals over time. So I think that some of the modifications that you can make even around the school improvement process that we have outlined would reduce the amount of teachers and principals having to resubmit a new plan every year. They are working on that same plan over a 3-year period to improve.

Mrs. BIGGERT. Thank you.

In Illinois I have been holding some roundtables and talking to teachers and talking to superintendents and then actually just the public, and one of the issues that people seem to really want to know more about or that they would like to have in their schools is the growth models. Is there somebody here who already—I know it is only, I think, two States—is there somebody here that has had to develop the growth model?

Ms. STRAUS. In Michigan we have not applied to use it. It has taken a very long time to develop the data system. As Mr. Jennings said, it is very complicated. And we have been working on this for a number of years and we are at the point now where we would be ready to apply for the growth model, and we would like to use it because we think it is a much better way of measuring. But it has taken us a long time. And I think a big State with so many students, we have 1.7 million students. Each individual record in

that system is much more difficult to develop than any of us thought would be the case.

Mrs. BIGGERT. Just one last thing.

The NAPE test. Have your schools been taking that test? I would just like to know what you think of it. Was it what you expected with the results versus the AYP?

Mr. MELMER. Yes. The NAPE exam is administered to grades 4 and 8 across the country. It is my understanding that all 50 States do participate in NAPE.

Mrs. BIGGERT. But it depends on the schools?

Mr. MELMER. Yes. Right. Correct. It is a sprinkling of students in every State. We certainly value the NAPE results in South Dakota. We think it gives us a third leg to the three-legged stool. We have our Dakota Step Test, which is our NCLB test. We have our ACT. In South Dakota, we give the ACT, which is sort of that regional postsecondary preparation exam. Then we look at the NAPE as being sort of that national comparison so that when we are at times criticized, saying "Your State test results look good, but how do you do at the national level?" we are prepared to come back and say, "Well, here at the NAPE, here are how our fourth and eighth grade students do." it is just that we have to be cautious about the NAPE, and that it is not designed to match State standards; it is certainly just a different type of exam, and so we look at it that way, but it certainly is one more measure that a State can look at to determine whether, in fact, their students are getting a good education.

Mrs. BIGGERT. Thank you.

Ms. JOHNSON. In terms of the value-added model, Dr. Sanders' model, it is used statewide. In addition to the status scores that the school districts get, each year we get graded in the core subject areas of the test based on value added, but what they do is they take a 3-year average, and so they look at the scores over a 3-year period, and they give you grades A to F, and that is reported on the State Web site per school.

The other part of the value added is they track and try to connect student results with teacher effects, and so each year the teachers get a teacher effects' score for how well the students—and again, this is over a 3-year period—have done. Principals have access to the data. One of the difficulties is that the teacher effects' score, based on State law, is not allowed to be used as part of the teacher performance review process.

We also use—we give or have in the past given the TSAT to all of our tenth graders, which is a preliminary SAT, and this year we are changing. We are giving the score tests at eighth grade, which is a preliminary ACT. We are giving planned tests at tenth, and then we are trying to get all of our students to take the ACT so that we have, in addition to the State assessment tools, some sense of how kids might do on another test.

Mrs. BIGGERT. Thank you.

Chairman KILDEE. Thank you, Mrs. Biggert.

The gentlelady from California, Ms. Woolsey.

Ms. WOOLSEY. Thank you, Mr. Chairman, for the good hearing. It is nice to see you, Jack.

Mr. JENNINGS. How are you?

Ms. WOOLSEY. Fine, thanks.

Jack—Mr. Jennings—I want you to answer this also and think about it, but I am going to start with Dr. Johnson.

Because we are hearing how many school districts are cutting back on history and geography and art and music and PE, I am really concerned that No Child Left Behind has led us to ignore the whole child; but it appears that you have been able to maintain a vibrant music program in Memphis at the same time as improving your AYP performance scores.

What can you tell us about how you have been able to sustain, and what can we learn from this important success?

Ms. JOHNSON. Well, let me just say quickly that I think this is an area where President Straus and I do agree very strongly on the importance of having music and arts programs for student engagement. Also, if you look at the standards that have been set by the corporate community, they want people who are creative, who are innovative, who can think outside the box, and we believe that the arts and music programs give those skills preparations that are necessary.

We use best practice, which is common planning time, and what we do is, in order to give teachers common planning time, we are employing people like music and art teachers to provide that release time so that those teachers, as a group of teachers who are working together with a group of students who are not doing well, can have the time to plan. And so we have not eliminated music and arts. In fact, we emphasize it as an important way to promote student engagement.

Ms. WOOLSEY. So, President Straus, would you just be saying, "Me, too," or do you want to add to that?

Ms. STRAUS. Yes, thank you.

We are very concerned about that. We think that the loss of creativity is really a great loss. And I was in China last year with a group of educators, and all they wanted to know is how do you teach creativity? And the restrictions that we put—and so much emphasis is put on testing and emphasizing the math and English, which is important, but it should not be to the exclusion of everything else, not only music and art and other arts, but what you raised about history and government and geography. I am a strong proponent—our whole board is—of social studies and the importance of civic education, and I have had the privilege of attending several of the congressional conferences on civic education, and I commend Congress for focusing on that because I think that is absolutely critical. It was one of the foundations of why we have public education, and I think that we put so much emphasis on the testing and on those major subjects that we do not have enough time for the others. But in Michigan, our own State accreditation system does test and does require social studies, and we do test in social studies as well. I know that is sort of contradictory, to test more, but we figure, if we do not test it, it is not going to get taught under the current system. So we are very concerned about that.

Ms. WOOLSEY. So, Mr. Jennings, you have some really important studies that you have referred to. Would those studies in any way indicate whether we are losing our whole child focus?

Mr. JENNINGS. Yes. In several weeks we are issuing another study where we ask a national sample of school districts how many minutes they spend on each curriculum subject and whether that has changed over the last number of years, and so we also ask them a number of other questions about instruction and curriculum, and we will issue that report towards the end of the month. We wanted to go beyond assertions to get data, and this is the same type of national sample of school districts. There are over 400 school districts. It is the same type of sample the U.S. Department of Education uses and others use, and we ask them very precise questions at the elementary level, the middle level, and the high school level, by subject area, and estimate the number of minutes. So we will have information in a short while.

Ms. WOOLSEY. Okay. We need that ASAP. Thank you very much. Thank you, Mr. Chairman.

Chairman KILDEE. Thank you.

The gentleman from New York, Mr. Kuhl. He passes.

The gentlelady from California, Mrs. Davis.

Mrs. DAVIS OF CALIFORNIA. Thank you, Mr. Chairman.

Thank you to all of you for being here as well.

I wanted to go to you, Mr. Jennings, to just quickly talk about data collection, and I think you mentioned the need to be doing this. And yet at the same time, school districts and States are really strapped to do that.

In California we have been pretty slow, actually, in data collection overall. So I am just wondering, how do we do that? Is it to require more of States that they need more resources to do that? What do you think ought to be done?

Mr. JENNINGS. Well, you know, every time somebody in elected office says you have to have more accountability, it transfers into a form that a local official has to fill out, and so we just have to recognize that. As national leaders and, I hope, State leaders, if you think of what you want to demand from people to get accountability, you have to think of the consequences. So I would hope that—and this is a perennial problem with government programs, but it is also a problem in private industry, too. I would hope that you would look at the law and figure out what you really need and what you do not need. I think you need test data, and I think you need test data all the way down the line, but you also need to spend some money to help school districts collect this data in the correct way. You need money to help States to make sure that they do it in the correct way, and you need some money for teachers so that they can understand the data to bring about improvement.

What is happening now with States is, with all this accountability, you know, testing in grades 3 through 8 and once in high school, some States have gone from higher-quality tests to lesser quality tests because they had to go from a few tests to a lot of tests, and they did not get the additional money they needed to have high-quality tests across the board. So you have to be aware of the fact that every time you ask for something, you have to think of the repercussions down the line.

Mrs. DAVIS OF CALIFORNIA. Do you think that designating that money only for tests is a wise——

Mr. JENNINGS. Congress has made a good start in that it has designated a pot of money to the States for tests. But what is happening now is, as a consequence of all this accountability testing, you are leading to a booming in what is called a "formative testing" because teachers cannot use this accountability testing. The test results do not arrive in time for them to use it to change education. So what they are doing is they are putting in more testing during the school year, which is not for accountability but for diagnostic purposes, so that they can understand where kids are as they proceed along, so that by the time they hit the accountability test they will be ready. Which means they are spending more money for testing, but this is for testing they think is useful rather than just for accountability.

So we have to think through the consequences of what we ask for, and once we ask for them as a Nation, we have to make sure that States and local school districts have the money to do it and that they do it in a way where they can use the information to improve education.

Mrs. DAVIS OF CALIFORNIA. Do any of you want to comment on that?

Ms. STRAUS. I mean one of the difficulties, too, is that I think in many cases, students are doing better on State tests than they are on NAPE, for example, and so there is a concern that there is a disconnect there.

Mrs. DAVIS OF CALIFORNIA. Mr. Secretary.

Mr. MELMER. And I would agree that you are probably going to see students do better on the State tests, oftentimes because it is matched directly to the State standards, and teachers know the standards and are teaching to those standards in the classroom versus the NAPE, which is more of a general exam and does not match up as well.

We do have a good model in place in our country right now in terms of addressing some of the things that Mr. Jennings mentioned. The NAPE program places a NAPE coordinator in every State, and that coordinator's job is to help facilitate the administration of that test and also to work with the national NAPE office on how the test works and all the recruitment and articulation of that test. We think CCSSO has always had discussions, and I think the U.S. Department is actually open to discussions about this topic where a data collection person could be placed in every State education agency, funded at least in part by the U.S. Department of Education, so that way, uniform training could go on; at least a consistency in language could take place, and we could begin to coordinate that effort at the national level rather than allowing every State to kind of have its own set of rules and regulations.

So we do see some potential solutions on the horizon. It is a matter of investing some dollars in that area. As Mr. Jennings said, if you expect it, then you have to sort of let the money follow that expectation.

Mrs. DAVIS OF CALIFORNIA. Right. All right.

Thank you, Mr. Chairman.

Mr. JENNINGS. Could I comment on the comparison between State test scores and NAPE?

The reason you have a disparity is for two reasons, principally. One is any test is a reflection of a curriculum or of standards. The NAPE test is a reflection of national standards, which, in a way, is a national curriculum, and State tests are a reflection of State standards or State curriculum. The two do not match up.

In Texas, for instance, they do not teach math in the way that is anticipated in the national math standards that are embedded in NAPE, and they get different results, therefore, because they are testing to something different. Texas has decided they want to test math "this way," but the national assessment says they are going to measure it "this way."

The second reason you have a difference in results is that no one child takes a full NAPE test. You have NAPE tests taken by different groups of children. There are no consequences.

Mrs. DAVIS OF CALIFORNIA. Do you see us ever aligning that?

Mr. JENNINGS. Well, there are bills in Congress—Senator Kennedy has one, Senator Dodd has one—to give funding to States to encourage them to move in the direction of adopting national standards as part of this reauthorization. That is a debate you are going to have, and that is what Mr. Finn was recommending, some consideration of national standards or an encouragement towards a national direction. But if you do not do that, you have to understand you will always have a discrepancy because of the curriculum matter, but also because of the motivation.

Kids know State tests count because teachers tell them, and there is all the pressure that is there to raise the State test scores. With NAPE, there is not that pressure, so kids do not put in the effort in NAPE that they put into State tests, and that is going to have an effect on results also.

Chairman KILDEE. Mr. Finn.

Mr. FINN. He left out the other big possible explanation for this discrepancy, which is that a lot of States have made it very easy to pass their State tests. And if you are seeing in a given State that the State says 70 percent of its fourth graders are proficient and NAPE says 27 percent of its fourth graders are proficient in that same subject in that same State, it might be because the States made it really easy to be defined as "proficient."

This is not necessarily a good thing for the people of that State. It might even be termed "misleading" for the people in that State to be told that their kids are proficient when, by national or world standards, they actually are not.

Mrs. DAVIS OF CALIFORNIA. Go ahead.

Chairman KILDEE. The gentleman from Illinois, Mr. Hare.

Mr. HARE. Thank you, Mr. Chairman, and my apologies for not being here sooner. If these questions have been asked or discussed, I apologize for reasking, but I am very interested in this.

One of the things I have heard universally back in my district from educators and from parents are the problems with our kids in special ed and the IDEA group. I guess I have a question for you, President Straus, and maybe also for you, Mr. Jennings, and maybe for the whole panel. And I think probably one of the biggest challenges that we have had on this committee is, you know, figuring out what to do and how we are going to change this and make it work.

How do we ensure that this group, especially that 2 percent that has cognitive abilities, is not left behind and held accountable while appropriately being tested? Is the problem that we just do not have enough data for the subgroup of students? So I am just sort of interested in, from your perspective, if you were advising us—which you are—you know, what can we do to make this work better for that group of students and educators?

Ms. STRAUS. Thank you. We are very concerned about that, because in Michigan we have special education from birth to age 26, and we want to keep that. But we also have people moving into our State with special needs children because we have a good program, so we have a higher percentage than many other States, and we think that the limit of 1 percent or 2 percent is not fair. It does not work right, and we would like to see that change.

That is one of the recommendations that we have, that students with disabilities should be assessed appropriately. We do not think they are now. A higher percentage, maybe, than 2 percent should be allowed if you can justify it and if you can really show that you have that many students. So that is one of the things that we are concerned about.

Mr. JENNINGS. If I could comment on that, too.

You have hit upon a very sensitive issue, and this is what teachers complain about throughout the country, and it is hurting No Child Left Behind because they are saying that No Child Left Behind, by holding these two groups of students to these standards, is making the whole goal not accomplishable. And so I would spend a considerable amount of time thinking through what you are going to do with these two groups. With children who are learning English—but let me start out by saying, with both groups, No Child Left Behind has led to much greater attention to their academic performance than ever before.

Children who are learning English are getting much greater attention in learning English and with academics than before. Children with disabilities are getting much greater attention. So what we have is a situation where there is a good being achieved—namely, these two groups are getting more attention—but they are being tested inappropriately, and teachers are complaining about the inappropriate testing. And so I think you are going to have to put much more flexibility in both of those areas.

One possibility with the English-learning students is that you combine two objectives—one is learning English, and the other is learning academic content—and you take kids who are new to the country or who do not know English, and you put most of the weight on their learning English. And then as they stay in school, you gradually shift the weight towards the academic content so you have a combined index. So that by the time you have gone several years, they have not only been measured on how they have learned English, but they have also been measured on the academic content. But you do not measure them on the academic content when they do not understand English and, therefore, cannot do well on the test. That is one possibility.

With children with disabilities, we had a meeting of the major organizations in Washington on the disability issue, and they told us there is no scientific basis for this 1 percent rule or this 2 per-

cent rule. These are just numbers that were chosen. I think in that area, you are going to have to give more flexibility so that individual children are given attention, so that there is more attention paid to the individual abilities of children to do well on tests, whether they should be held to the same standards or whether they should be tested the same way, and figure out some way so that it is more personalized than it is today.

Otherwise, what you are going to have with this 1 percent/2 percent rule are States just putting enough children in to amount to 1 percent or amount to 2 percent, and they will be called "2 percent children" and "1 percent children." they are an arbitrary number. You have to pay attention to the individual needs of these kids much more, I think.

Mr. HARE. Dr. Johnson, I was just wondering if you have any thoughts beyond any of the other panelists, because I know this is very important to me.

Ms. JOHNSON. Well, I do think that, including special needs and ELL, students have improved significantly the focus—and that is an important focus—but I will say that in urban districts in particular, the higher number and higher percentage of students with special needs usually means that in the urban areas, we reach the threshold cutoff, and in some smaller districts they do not reach it.

So it is more likely you will see that urban districts get identified more quickly in the way it is set up and not making adequate yearly progress for special needs students.

I know in our county, even though we are not the only district in our county, as for the students with special needs, particularly the low-incidence population—if they are blind or deaf—we serve them all because some of the smaller districts cannot.

What you do not want as an unintended consequence is, you do not want districts to refuse or act like they do not want to serve these kids because they do not want to be on the list when you can have some efficiency in serving deaf children in a more concentrated way than if every little district had to do that that is in close proximity.

So I think it is important for you to be very careful about the sanctions when these kids deserve a good education, and you do not want to incent people not to be accepting of the diversity of students' needs that exist in our community.

Mr. HARE. Thank you.

Thank you, Mr. Chairman.

Chairman KILDEE. The gentleman from New Jersey, Mr. Payne.

Mr. PAYNE. Thank you very much, Mr. Chairman.

Let's see. President Straus, many educators have called for the use of multiple measures to assess school districts and schools under No Child Left Behind.

Can you discuss additional factors that you believe ought to be taken into consideration in determining a school's effectiveness, in addition to the test currently used?

Ms. STRAUS. Thank you very much for that question. We think it is not right to measure schools and students based on one test on a given day, and we think there are other things that should be taken into account.

In our own accreditation system, we look at what else is being taught, what is being taught in the school. Are they providing social studies? Are they providing arts? Are they looking at other—are they looking at the other subjects that are being considered? What is the graduation rate? What is the dropout rate? Is there parental involvement? What is the relationship with the community?

All of these things go into measuring a school. When you walk into a school, you do not know whether—you can tell whether the school is a good school when you walk into it, and partly it is because of the parental involvement; it is the quality of the teachers; it is all of these things that should be taken into account when you are measuring how good a school is. And it is not just a given test on a given day.

Thank you.

Mr. PAYNE. Would anyone else like to respond?

I am trying to think back—and maybe you could think back, too—as to how we were assessed. I mean we had regular tests, but I do not remember too many of these high-stakes tests in the fourth grade or in the sixth grade or—you know, maybe out of elementary school, perhaps, or something like that. And all of my classmates seemed to have learned—I mean I am excluding me, perhaps—but they did well without these high-stakes tests.

So I think there needs to be—I mean we are all for accountability, but it seems that there are other ways to measure achievement than that.

But I just have another quick question for—Dr. Johnson, is it? Just rushing through everyone's testimony here quickly, I did notice that you talked about flexibility and so forth, and you talked about recruitment support, employment of staff in high-poverty districts. And this is really something that I sound like a broken record with, about the opportunity to learn; that youngsters need the opportunity to learn, and they are not given the opportunity to learn when they have inexperienced teachers. And as we know, the system just tends to put good teachers or better teachers into a situation they feel is better because that's what, you know, seniority does.

So do you have any thoughts on how we could have incentives, or, to kind of turn that around so that the schools in most need do not continually get the least experienced teachers?

Ms. JOHNSON. Congressman, what we have done in Memphis—and I can speak to that—is we have fresh-started, reconstituted, 12 schools. We negotiated a separate agreement with our collective bargaining group so that we could hire people totally out of seniority order. We created two incentives, a front-end incentive and a back-end incentive.

The front-end incentive was to give people 2 more weeks of pay, but they had to participate in the professional development that we design to work in that school, and the principal was able to hire anybody whom he wanted to hire. We took very careful consideration of who the principal was, clearly.

Then the back-end incentive was if they made progress, they could get anywhere from $1,000 to $3,000 in additional compensation. Now, this was very targeted, in a pilot way, at the schools that were in high priority that we had reconstituted.

I do think incentives matter. I do not think that teachers necessarily go into education to necessarily—they have never thought about these kinds of reward systems. But I think that it is important for us to rethink the connection between student performance and giving teachers rewards and recognition, especially as it relates to making sure that we have teacher induction and mentoring programs. The kind of flexibility that I think we are asking for, both in terms of transferability and in other ways within the No Child Left Behind, would give us the flexibility to spend dollars in ways that I think would create incentives and support for teachers in the profession.

Mr. PAYNE. Thank you. Thank you very much.

Chairman KILDEE. I was talking to Mrs. Biggert here. I will maybe throw one more question out here, and all of you can answer it.

I think both Dr. Finn and Jack, you have touched on it and talked about it, but NAPE and many or most State tests are testing to different standards or to a different curriculum. Should we in this reauthorization do some things to encourage States to bring their tests so they are testing against a more common curriculum or standard?

We will start with you.

Mr. FINN. "Yes, sir" is the answer.

The variability in State standards has not led to good standards in America. My foundation reviews State academic standards approximately every 5 years and, to be perfectly honest, most of them are thoroughly mediocre in terms of what they actually expect kids to learn. Then you add to that the problem of variable State expectations on their own tests, and the result is something akin to chaos that a big, modern, competitive country should not tolerate. And while I doubt that you could or should even create a compulsory national standard or test, I think to have a voluntary version that you incentivize States to join in with, leaving them the option of staying out if they would rather, would be a very, very, very important reform in this next round of NCLB.

Chairman KILDEE. Thank you very much.

Jack, and then I will let the rest of you give your views.

Mr. JENNINGS. My view is I do not think we are going to get a national curriculum in this country, and I do not think the Congress is going to be able to mandate a national curriculum or to mandate national standards.

If you want the States to have higher standards, I think you are going to have to make it attractive to them by giving them financial incentives or by giving them regulatory relief. And you, at a minimum, will have to help them pay for rethinking their standards in order to align them with national standards, and that is what Senator Kennedy's bill does; that is what Senator Dodd's bill does.

Chairman KILDEE. Ms. Johnson.

Ms. JOHNSON. Yes.

Every one of our students will take a college entrance exam, either the ACT or the SAT, and if they go to a community college, they will take some kind of entrance exam that will place them either in a remedial or in a regular English course. So I think that

from that perspective—that is the perspective I come with—not to prepare them to be able to enter college and be successful and have some notion that their diploma means something is problematic.

So, while I do not think that I am talking about a national curriculum, I do think that a lot of the States—for example, when the National Council of Teachers of Mathematics defined their standards, a lot of States modified their standards to be in alignment with them. And so I think we have to look at these things.

Now, the Council of the Great City Schools has nine districts that are testing a larger sampling of NAPE so that there are more kids in those urban districts—New York, Boston. So I think that supporting some of those pilots where we would get a greater sampling, we would be able to do some comparability across State lines. But the kids are still going to have to get into college or postsecondary education, and there are not a lot of hoops that are different for them to get there.

Chairman KILDEE. You know, Jack, you were here, I think, in my first term that I was here. We established——

Mr. JENNINGS. I welcomed you.

Chairman KILDEE [continuing]. The U.S. Department of Education, and I think we put language in there actually saying that the Federal Government would not establish a national curriculum.

Mr. JENNINGS. That is correct.

Chairman KILDEE. That language is still in there.

Mr. JENNINGS. And it was repeated in the reauthorization that is the No Child Left Behind Act. Federal officials cannot dictate curriculum at the local level.

Chairman KILDEE. When you set standards, you are kind of moving towards touching curriculum, though, are you not?

Mr. JENNINGS. What is happening now is we have an extremely decentralized system of public education. Almost every other industrialized country has a national curriculum or national standards, and they can anchor what they do around what is being taught. In the absence of a national curriculum or a national standard here, we have a fad of the day, local school-based management or something else, and then we twirl around, and we wonder why we do not have increases in achievement. It is because we are not paying enough attention to the curriculum, to what is being taught, to what should be behind the tests.

If we are not going to have a national curriculum, what is happening now with the standards-based reform is that the States are gradually moving toward State curriculums. Some States are more up front about this than others. In North Carolina they are up front about this, that they have a State curriculum. In Maryland they have a voluntary State curriculum. So States are gradually saying, this is what we expect kids to know.

You are probably best off encouraging States to move in that direction because they have enough trouble dealing with local school districts, even going in that direction, rather than talking about a national curriculum. But if you want more uniformity among the States with curriculum, you are going to have to do it through some incentive basis.

Mr. Kildee, can I answer Mr. Payne? Let me leave one word or a couple words with you about teachers.

I know everybody is worried about the quality of teachers in the poorest schools and in the schools with the highest numbers of children of color. And it is a serious national problem, but the tendency of the Congress will be to take one little solution and enact it, like a mentoring program for new teachers, and think that that is the answer. That is not the answer.

We convened all of the major education organizations that deal in this area, and we asked them what should be done with this problem. The answer—and it was very strongly supported—was that it has to be a comprehensive approach, and it really should be done from a State level dealing with local school districts. And the comprehensive approach should deal with the pretraining of teachers so that they know what they are going to face when they meet kids who are culturally different. It should deal with higher pay for teachers who are going into schools that are very challenging. It should deal with mentoring programs that are of high quality. It should deal with changing the conditions of education for new teachers or the conditions of teaching so that they have more time to prepare. It should deal with a variety of different things, but if you cherry-pick and just do one little program, you are not going to add to this solution.

It has to be a comprehensive approach, dealing with everything—conditions of education, pretraining, mentoring, pay. It has to be a comprehensive approach, and it has to be approached, I think, from the State level, getting beyond individual school districts so that they can recruit more broadly. But I would hope you would pay a lot of attention to this issue, because the quality of teaching frequently determines how well kids do in school.

Chairman KILDEE. I want to thank all of you.

This has been an excellent hearing, a very good hearing, and I think we have gotten some solid information, some solid views that will help us, and this will become part of the body of knowledge that we will use in reauthorizing this bill. So I want to thank all of you.

Again, as previously ordered, members will have 7 additional days, and any member who wishes to submit follow-up questions in writing to the witnesses should coordinate with the majority staff within the requisite time.

Without objection, this hearing is adjourned.

[Whereupon, at 4:27 p.m., the subcommittee was adjourned.]

○