AUTHENTICATED
U.S. GOVERNMENT
INFORMATION
GPO

S. HRG. 111–1118

ESEA REAUTHORIZATION: STANDARDS AND ASSESSMENTS

HEARING

OF THE

COMMITTEE ON HEALTH, EDUCATION, LABOR, AND PENSIONS

UNITED STATES SENATE

ONE HUNDRED ELEVENTH CONGRESS

SECOND SESSION

ON

EXAMINING ELEMENTARY AND SECONDARY EDUCATION ACT (ESEA) REAUTHORIZATION, FOCUSING ON STANDARDS AND ASSESSMENTS

APRIL 28, 2010

Printed for the use of the Committee on Health, Education, Labor, and Pensions

Available via the World Wide Web: http://www.gpo.gov/fdsys/

U.S. GOVERNMENT PRINTING OFFICE

56–288 PDF WASHINGTON : 2012

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

VerDate Nov 24 2008 12:34 Mar 08, 2012 Jkt 000000 PO 00000 Frm 00001 Fmt 5011 Sfmt 5011 S:\DOCS\56288.TXT DENISE

COMMITTEE ON HEALTH, EDUCATION, LABOR, AND PENSIONS

TOM HARKIN, Iowa, *Chairman*

CHRISTOPHER J. DODD, Connecticut
BARBARA A. MIKULSKI, Maryland
JEFF BINGAMAN, New Mexico
PATTY MURRAY, Washington
JACK REED, Rhode Island
BERNARD SANDERS (I), Vermont
SHERROD BROWN, Ohio
ROBERT P. CASEY, JR., Pennsylvania
KAY R. HAGAN, North Carolina
JEFF MERKLEY, Oregon
AL FRANKEN, Minnesota
MICHAEL F. BENNET, Colorado

MICHAEL B. ENZI, Wyoming
JUDD GREGG, New Hampshire
LAMAR ALEXANDER, Tennessee
RICHARD BURR, North Carolina
JOHNNY ISAKSON, Georgia
JOHN McCAIN, Arizona
ORRIN G. HATCH, Utah
LISA MURKOWSKI, Alaska
TOM COBURN, M.D., Oklahoma
PAT ROBERTS, Kansas

DANIEL SMITH, *Staff Director*
FRANK MACCHIAROLA, *Republican Staff Director and Chief Counsel*

VerDate Nov 24 2008 12:34 Mar 08, 2012 Jkt 000000 PO 00000 Frm 00002 Fmt 0486 Sfmt 0486 S:\DOCS\56288.TXT DENISE

C O N T E N T S

STATEMENTS

WEDNESDAY, APRIL 28, 2010

	Page
Harkin, Hon. Tom, Chairman, Committee on Health, Education, Labor, and Pensions, opening statement	1
Enzi, Hon. Michael B., a U.S. Senator from the State of Wyoming, opening statement	2
Paine, Steven L., Ph.D., Superintendent of Schools, West Virginia Department of Education, Charleston, WV	5
Prepared statement	7
Schmeiser, Cynthia B., Ph.D., President, Education Division, Act National Office, Iowa City, IA	10
Prepared statement	12
Phillips, Gary, Ph.D., Vice President, American Institutes for Research, Washington, DC	17
Prepared statement	19
Rivera, Charlene, Ed.D., Executive Director, George Washington University Center for Equity and Excellence in Education, Alexandria, VA	25
Prepared statement	27
Thurlow, Martha, Ph.D., Director, National Center on Educational Outcomes, Minneapolis, MN	32
Prepared statement	33
Murray, Hon. Patty, a U.S. Senator from the State of Washington	47
Alexander, Hon. Lamar, a U.S. Senator from the State of Tennessee	48
Franken, Hon. Al, a U.S. Senator from the State of Minnesota	50
Isakson, Hon. Johnny, a U.S. Senator from the State of Georgia	51
Bennet, Hon. Michael F., a U.S. Senator from the State of Colorado	53
Hagan, Hon. Kay R., a U.S. Senator from the State of North Carolina	56
Casey, Hon. Robert P., Jr., a U.S. Senator from the State of Pennsylvania	58

VerDate Nov 24 2008 12:34 Mar 08, 2012 Jkt 000000 PO 00000 Frm 00003 Fmt 0486 Sfmt 0486 S:\DOCS\56288.TXT DENISE

ESEA REAUTHORIZATION: STANDARDS AND ASSESSMENTS

WEDNESDAY, APRIL 28, 2010

U.S. SENATE,
COMMITTEE ON HEALTH, EDUCATION, LABOR, AND PENSIONS,
Washington, DC.

The committee met, pursuant to notice, at 2:05 p.m. in Room SD–430, Dirksen Senate Office Building, Hon. Tom Harkin, chairman of the committee, presiding.

Present: Senators Harkin, Bingaman, Murray, Casey, Hagan, Merkley, Franken, Bennet, Enzi, Alexander, and Isakson.

OPENING STATEMENT OF SENATOR HARKIN

The CHAIRMAN. The Senate Committee on Health, Education, Labor, and Pensions will please come to order.

Today's hearing will focus on the important role that standards and assessments play in our education system.

In our previous ESEA hearings, a variety of experts have impressed on us the importance of our country developing a world-class education system that prepares our students to be successful after high school graduation. In order to do this, it is vital that we have a clear understanding of what students need to learn and develop ways to accurately assess their progress to determine what they are learning and where they need additional help.

Nearly 30 years ago, the landmark report, "A Nation at Risk," highlighted the need for rigorous standards in our country's schools. About a decade later, the Nation's Governors heeded that call with the Charlottesville Summit, and the Federal Government supported States' efforts to develop standards by passing Goals 2000 and the Improving America's Schools Act, which was the 1994 version of ESEA.

At the beginning of the last decade, we took the next steps by requiring that all students within a given State be held to the same high standards. These standards helped to end a two-tiered system that meant lower expectations for disadvantaged students. However, the standards did not ensure that students were being prepared for success after high school graduation.

In Iowa, for example, over 80 percent of high school graduates plan to pursue training or college after high school. Yet, too often, they are unprepared to meet the challenges of post-secondary education. Experts estimate that nearly 60 percent of students entering post-secondary schools need to take remedial courses to catch up to college-level coursework.

The Alliance for Excellent Education has estimated that this need for remediation costs our Nation at least $3.7 billion every

VerDate Nov 24 2008 12:34 Mar 08, 2012 Jkt 000000 PO 00000 Frm 00005 Fmt 6633 Sfmt 6633 S:\DOCS\56288.TXT DENISE

year. The problem is also evident in the workforce. A recent study estimated that over 50 percent of high school graduates do not have the skills to do their job, compared to less than 20 percent of college graduates.

While the adoption of State standards was no small achievement in No Child Left Behind, it is clear that, as we reauthorize this bill, serious improvements are necessary. We must ensure that the standards that States set are not false benchmarks but translate into success, whether students chose to go to college or enter a career.

I might also add that there are important civil rights and equity questions at play here also. Professor Goodwin Liu of Boalt Hall Law School at Berkeley published a paper showing that those States with the highest minority and low-income populations also tend to have the lowest standards.

Finally, the obvious issues of teacher preparation and economies of scale are central to this conversation. How can schools of education properly prepare teachers to teach to standards if those standards may be significantly different in the State where the teacher ends up teaching after graduation?

We do not have a mandate that says if a teacher goes to the University of Northern Iowa and takes a course in education to become a teacher that that person has to stay in Iowa all their life. They may go to Minnesota, and a lot of them do, quite frankly.

I applaud the leadership of the chiefs and the Governors and their partners in developing this Common Core, and I look forward to hearing more about this. However, along with setting high achievement goals, we must also develop the ability to measure whether or not students are meeting those goals.

Because of NCLB's testing requirements, we know more about which students are achieving and which need more assistance and support. Teachers need to know this, too. However, in many cases, that measurement is being done through low-quality tests that don't measure the range of skills and knowledge that we value.

Technological advancements have made it possible to adapt questions during a test to better show the depth of a student's knowledge of the subject or to electronically score short-answer or essay questions, not just multiple choice.

In this reauthorization, it is critical that we redouble our commitment to ensuring that students will graduate ready to meet the challenges of college and the workplace. As we have heard time and time again, our economic success in the next century is directly tied to our ability to have a highly educated, highly skilled workforce.

I look forward to hearing from our panelists today because adopting high-quality standards and assessments is an important step to that end. I thank all of them for being here. After Senator Enzi makes his opening statement, I will introduce the panel, and we can hear from the panel, and we will open it up for discussion.

With that, I recognize Senator Enzi.

OPENING STATEMENT OF SENATOR ENZI

Senator ENZI. Thank you, Mr. Chairman.

I just wanted to note that all of them don't go to Minnesota. Some of them come to Wyoming. That way, they don't have to learn a new accent.

[Laughter.]

I do want to thank you for continuing this series of hearings on the reauthorization of the Elementary and Secondary Education Act with the important issues of standards and assessments. The witnesses before us today have provided some excellent written testimony and will provide insight and information that will be very helpful in our work to reauthorize ESEA.

I want to start by applauding the work of the National Governors Association and the Council of Chief State School Officers in the development of the Common Core standards. This effort was appropriately led, developed, and should be continued by the States.

I have said for many years that students need to be provided with knowledge and skills they need to be successful in college and the workforce. The Common Core standards, developed by the States, if implemented and adopted properly, could finally move our country in that direction.

However, the Federal Government should stay out of the way of these efforts. As we work on the reauthorization of ESEA, we should find ways to assist States, not require or coerce them with this difficult, but important, work. The development and adoption of these standards by the States are just the first steps in a very long process.

Once adopted, States are going to have to implement new assessments and curriculum aligned to these standards. This process will take time, which may be longer for some States than others, but it has the potential to save money in the cost of test validation. Part of the process is making sure that the impact on teachers in the classroom is positive and that they are given the training and support they need to teach students.

I am pleased that many of you will also discuss assessments. While considering the changes to ESEA, we need to maintain the high standard of including all students in single State-wide accountability systems. The growth of individual students and collectively among a group of students in reducing the achievement gaps between higher-achieving and lower-achieving students has been well-documented. We cannot stop moving in this direction now as we continue to prepare our students, rich or poor, with or without disabilities, or English language learners for post-secondary education or employment in the global economy.

However, when assessing these students, we need to be sure they are taking the assessment that best measures their ability. Therefore, students need access to the necessary and proper accommodations and other supports they need to accurately reflect their true ability and capability.

The work being done by States on standards has spearheaded significant discussion among the next generation of State systems of assessment. In my travels across Wyoming, I hear over and over that the static model used by many States under No Child Left Behind needs to be changed to allow for growth models in all the States.

I am particularly pleased that these new assessments will be better aligned to allow for better measurements of student growth from year to year. It is important to maintain regular assessments that summarize the development of students so that we know how a student has done over the course of each year.

It is also important to support State systems of assessment that would include various assessment models, many of which could be used by teachers to better inform the work they do in the classroom. I also believe that these new assessments can do a better job of measuring higher-order thinking skills and the 21st century skills that business leaders need in their workforce.

All of these changes will also have a huge impact on the data that we report, collect, and use. As we work through these changes, we must remember that it is important to measure what we value instead of valuing what we measure. States and school districts have developed data systems, but it is still unclear how much of that data is accessible by teachers to really have an impact on their work in the classroom.

Elementary and secondary education in this country is undergoing some new and exciting changes. Our work on the reauthorization of ESEA must be done carefully and deliberately to foster and support the changes. NCLB is often criticized for its unintended consequences. If we are not thoughtful and instead work quickly because we are trying to meet artificial deadlines, we could wind up being criticized even more than we are now.

I want to welcome all the witnesses and thank them for being with us today to share their knowledge and expertise. I know that we won't have a chance to ask you all the questions that we need to, based on the testimony that I have already read. I hope that you realized you have volunteered to answer written questions that we might have afterwards.

I look forward to learning more from each of you in the efforts you have undertaken in the areas of standards and assessments.

Thank you, Mr. Chairman.

The CHAIRMAN. Thank you, Senator Enzi.

I just want to reassure you and also my friend from Minnesota, Senator Franken, that because so many Minnesotans come south to Iowa for the winter——

[Laughter.]

We like to be welcoming, we make sure that all of our kids in Iowa are taught to say, "Ya, you betcha."

[Laughter.]

That is Minnesotan. Just want to reassure you that we do know how to speak that language.

Senator FRANKEN. Oh, thanks a lot.

[Laughter.]

The CHAIRMAN. Ya, you betcha.

[Laughter.]

Now I will thank our panel for being here. I will introduce you all. Then we will just go in order.

First, we have Dr. Steven Paine, West Virginia's 25th State superintendent of schools. Under the leadership of Dr. Paine, West Virginia has been internationally and nationally recognized for its 21st century learning program, its pre-K programs, school leader-

ship development programs, reading initiatives, and teacher quality efforts. Dr. Paine was recently elected president of the board of the Council of Chief State School Officers.

Next, from my home State of Iowa, someone who did not go to Minnesota, Dr. Cindy Schmeiser. As the Education Division president at ACT, Dr. Schmeiser is responsible for leading and coordinating the research, development, and client support of all assessment instruments associated with ACT's educational programs. She obtained her Master's and Doctorate degrees in educational measurement and statistics from the University of Iowa.

Next, we are joined by Dr. Gary Phillips. Dr. Phillips is a vice president at American Institutes for Research. He has served as the commissioner of the National Center for Education Statistics from 1999 until 2002. He is internationally known for his expertise in large-scale assessments and complex surveys.

Next, we welcome Dr. Charlene Rivera, who directs the George Washington University Center for Equity and Excellence in Education. As a nationally recognized education researcher, Dr. Rivera is known for work that addresses English language learners. Her research focuses on State assessment policies and practices, accommodations, accountability, national standards, program evaluation and reporting, and leadership development.

Finally, we are grateful to have Dr. Martha Thurlow, who is the director of the National Center on Educational Outcomes. In this position, she addresses the implications of contemporary policy and practice for students with disabilities, including national and Statewide assessment policies and practices, standard-setting efforts, and graduation requirements.

Dr. Thurlow has conducted research for the past 35 years in a variety of areas, including assessment and decisionmaking, learning disabilities, early childhood education, dropout prevention, effective classroom instruction, and integration of students with disabilities in general education settings.

We have a very distinguished panel to talk to us about standards and assessments. We welcome you here. All your statements will be made a part of the record in their entirety.

I would appreciate it if you could take 5 minutes or thereabouts to just give us a summary of it, and then we can engage you in conversation.

Dr. Paine, welcome, and please proceed.

STATEMENT OF STEVEN L. PAINE, Ph.D., SUPERINTENDENT OF SCHOOLS, WEST VIRGINIA DEPARTMENT OF EDUCATION, CHARLESTON, WV

Mr. PAINE. Thank you, Mr. Chairman.

Chairman Harkin, Ranking Member Enzi, members of the committee, thank you for inviting me to speak today on behalf of the Council of Chief State School Officers and the National Governors Association regarding Common Core standards and that initiative and the State cooperation to develop high-quality assessments that is related to that project.

My name is Steve Paine. I am the State superintendent of schools in West Virginia and the current president of CCSSO.

And I just might add, Senator Harkin, Senator Enzi, in your opening remarks, you are spot on to the issues that exist around the Common Core standards and the assessments that are related. I am so appreciative that you understand the issues so clearly. Thank you.

As the committee continues to examine the Elementary and Secondary Education Act, I appreciate this opportunity to talk about the States' extraordinary leadership in a voluntary effort to ensure that all students are held to college- and career-ready standards. The State-led Common Core development effort will conclude next month with the release of the final grade-by-grade standards in English language arts and mathematics.

This work is representative of States' commitment to leading the way on education reform. Given this bold State action, we hope the Federal Government will respond in kind by ensuring that the updated ESEA supports a new State-Federal partnership, a partnership that provides States with greater authority to innovate and appropriate incentives and supports to help them not only implement college- and career-ready standards, but also improve teacher and leader effectiveness, strengthen longitudinal and instructional data systems, and turn around low-performing schools.

CCSSO and NGA launched the voluntary State-led Common Core standards initiative to provide a coherent foundation for ensuring that all students leave high school ready for college and career. The 48 States, including West Virginia—and I might add that we are one of those States that is a high-minority State.

We are very proud of the fact that we have taken this work of Common Core standards seriously, and our own new standards now receive top billing in most recent quality counts report in Education Week for being rigorous and having a very rigorous standards assessment accountability system. So it can be done, and we have rolled up our sleeves to say that we are going to accept that challenge.

As we move forward, there were five core principles that permeated the development of the Common Core standards. It was determined that the common standards must be, No. 1, fewer, clearer, and higher. A mantra that we learned from some of those higher performing nations out there, such as Singapore and Taiwan and others.

Second, that they needed to be, in fact, internationally benchmarked to the curricula of those highest-performing nations.

Third, that they include rigorous content knowledge, along with those skills that you have identified in your introduction that are so critical to business and into the private sector today in our global world.

Fourth, they must be evidence and research-based. And finally, that they prepare students for college and career. That is an absolute must.

As the development phase concludes, State adoption and implementation of the Common Core will help to ensure that all students are called upon to satisfy college- and career-ready standards and will enable fair and accurate performance comparisons—which is a key point of this, State comparability—between States, while

catalyzing and enabling unprecedented State collaboration to address the Nation's most pressing educational challenges.

Several States are in the process of adopting the Common Core. Kentucky led the way, and they were the first State. I am proud to say our State will adopt the Common Core next month, and I am sure other States will follow suit as the deadline for adoption approaches us.

Participating States developed the Common Core standards in two phases with the support of leading standards experts who collaborated with a range of interested stakeholders from across the country. The very transparent development process included numerous opportunities for public comment and benefited from constructive feedback provided by individual schoolteachers—very importantly schoolteachers—and leaders, national education organizations, higher education representatives, civil rights groups, and other interested parties and individuals.

The initiative's phase one work concluded in the fall of 2009 when CCSSO and NGA published Common Core and career-readiness standards, illustrating what students should know at the end of high school. Since that time, the initiative's second phase of work has focused on back-mapping the college- and career-ready standards on a grade-by-grade basis for kindergarten through Grade 12.

Let me just say in conclusion, if I might, that we certainly are excited about the opportunities that have been availed with the advent of $350 million in Race to the Top for innovative assessments to assess this Common Core, the full scope and—the full scope and range of the Common Core standards.

I also deeply appreciated in your opening remarks your acknowledgment that that certainly includes a summative standardized test, but must include other measures of progress, particularly those that are employed by teachers in classrooms as we look at multiple measures of how we assess progress.

I certainly appreciate the time that you have provided to me today to offer testimony and look forward to any questions that you might have for me after the testimonials from our other experts.

Thank you, sir.

[The prepared statement of Mr. Paine follows:]

PREPARED STATEMENT OF STEVEN L. PAINE, PH.D.

Chairman Harkin, Ranking Member Enzi, members of the committee, thank you for inviting me to speak today about the Council of Chief State School Officers (CCSSO) and National Governors Association (NGA) common core standards initiative and State cooperation to develop related high quality assessments. My name is Steve Paine, I am the State Superintendent of Schools in West Virginia and the current President of CCSSO.

As the committee continues to examine the Elementary and Secondary Education Act, I appreciate this opportunity to talk about the States' extraordinary leadership in a voluntary effort to ensure that all students are held to college and career-ready standards. The State-led common core development effort will conclude next month with the release of the final grade-by-grade standards in English Language Arts and Mathematics. This work is representative of States' commitment to leading the way on education reform. Given this bold State action, we hope the Federal Government will respond in kind by ensuring that the updated ESEA supports a new State-Federal partnership. A partnership that provides States with greater authority to innovate and appropriate incentives and supports to help them not only implement college and career-ready standards, but also improve teacher and leader effectiveness,

strengthen longitudinal and instructional data systems and turnaround low-performing schools.

SUMMARY

CCSSO and the NGA launched the voluntary State-led common core standards initiative to provide a coherent foundation for ensuring that all students leave high school ready for college and career. The 48 States, including West Virginia, two territories and the District of Columbia who worked collectively to develop the common core standards in English language arts and mathematics were guided by several core principles. The common standards must be: (1) higher, clearer and fewer, (2) internationally benchmarked, (3) include content knowledge and skills; (4) evidence and research-based; and (5) prepare students for college and career. As the development phase concludes, State adoption and implementation of the common core will help to ensure that all students are called upon to satisfy college and career-ready standards and will enable fair and accurate performance comparisons between States, while catalyzing and enabling unprecedented State collaboration to address the Nation's most pressing educational challenges.

Participating States developed the common core standards in two phases with the support of leading, standards experts, who collaborated with a range of interested stakeholders from across the country. The transparent development process, included numerous opportunities for public comment and benefited from constructive feedback provided by individual school teachers and leaders, national education organizations, higher education representatives, civil rights groups, and other interested parties and individuals. The initiative's phase one work concluded in the fall of 2009 when CCSSO and NGA published common college and career readiness standards, illustrating what students should know at the end of high school. Since that time, the initiative's second phase of work has focused on back-mapping the college and career-ready standards on a grade-by-grade basis for Kindergarten through Grade 12.

As the standards development work draws to a close and participating States begin the voluntary adoption process, several exciting common State assessment collaboratives are beginning to take shape, including a group co-led by West Virginia. State cooperation to develop common, high quality assessments is possible because of the common core standards initiative and will be furthered by the $350 million Race to the Top Assessment competition. These advanced assessment systems will measure student knowledge and skills against the full range of college and career-ready standards and will represent the next generation of summative, formative and interim assessments, which will significantly improve teaching and learning in the classroom by providing unprecedented insights into student's status and growth.

Even as this important State-led standards and assessment work continues, we are pleased to have this opportunity to make recommendations to Congress about how the Elementary and Secondary Education Act might be updated to support State-led education reform. The Nation's chief State school officers believe the new ESEA should continue and expand the Federal Government's strong commitment to supporting State assessment development, support movement toward voluntary State college and work-ready standards, and fund the creation of aligned, and enhanced assessments systems. New instrumentation should be fully representative of the richness of standards and allow for students who learn at different rates.

COMMON CORE STANDARDS DEVELOPMENT PRINCIPLES AND TRANSPARENCY

As mentioned earlier, even before State development of the common standards started in early 2009, the initiative was driven by four fundamental principles designed to ensure the integrity and quality of the standards.

First, the common standards are higher, clearer and fewer. Each design element was crucial to the development process. Higher standards raise the bar to prepare students for international competitiveness. Being committed to higher standards ensures that no State would lower its standards by adopting the common core. Clearer standards allow parents, students, and teachers to understand exactly what is expected of students as they advance through the system. Fewer standards allow teachers to more deeply focus on topics. One challenge that State leaders consistently hear from educators is that current standards are too numerous to cover in the school year. To overcome this challenge, we raised the bar and focused the standards to maximize student learning.

Second, the common standards are internationally benchmarked. American students are entering a global economy that requires competition with students from around the world. Through States' development of the common standards, we evalu-

ated other high achieving countries' standards to ensure that the common core represented world class standards. As a result, the new standards will prepare American students to be internationally competitive when they leave the Nation's public schools.

Third, the common standards development process was informed by evidence and cutting edge research. Historically, standards were often based largely on personal judgment. By allowing personal judgment to determine what concepts are in or out of standards, the process often became a negotiation, rather than a reflection on what evidence and research tells us about the connection between K–12 experiences and success in higher education and promising careers.

Lastly, the common standards are aligned with college and work expectations. By preparing all students to be both college and career-ready, all students are able to compete in their post-secondary education and/or career choice. Preparing all students to be college- and career-ready is absolutely critical to the long-term success of the country. By providing a set of expectations that are clear to students, parents and educators about what it takes to be college and career-ready, the States have taken a major step forward in producing students who are ready for later success.

COMMON CORE STANDARDS DEVELOPMENT PROCESS

CCSSO and NGA committed to participating States, territories and the District of Columbia that the standards development process would be open and transparent. In April 2009, over 40 States met to discuss the possibility of creating common core standards in English language arts and mathematics. Following this meeting, 48 States formally agreed to join the common core standards development effort and begin a two-phase process. Phase one: develop college and career readiness standards. Phase two: create college and career standards through K–12, grade-by-grade by Spring 2010. Using experts and practitioners from across the Nation and throughout the world, the States completed the college and career readiness standards in September 2009. The standards were reviewed by States, the public, and a range of national organizations and outside experts. Based upon the college and career-readiness standards completed late last year, participating States and the expert development team immediately began development of the grade-by-grade K–12 standards, successfully releasing the standards for public comment in March 2010. Public comments were due on April 2 and the final K–12 expectations will be released next month, allowing States to begin the adoption and implementation process.

States are responsible for demonstrating that within 3 years they have fully implemented the standards by developing instructional supports and aligning assessments. Kentucky has already formally adopted the common core and we expect a significant number of States to follow Kentucky's lead later this year.

BENEFITS OF COMMON STANDARDS: STUDENTS, PARENTS, AND TEACHERS

Common standards are a positive development for all students. The standards will help equip students with the knowledge and skills needed to succeed in college and careers. The new standards will also set high expectations for learning across the Nation, ensuring that all students must meet a high bar regardless of where they live. The standards will allow students to more easily transition from one State to another without losing valuable learning time adjusting to different standards. Given the mobility of the student population in the United States, common standards are essential. Also, higher, clearer, and fewer standards makes the student's responsibilities clear, so that they can take charge of their own learning.

For parents and other caregivers, common standards will delineate exactly what their student needs to know and be able to do at each educational stage. With clearer and fewer standards, parents will be better positioned to facilitate conversations with their child's teachers about what they should be learning and how they can reach their goals creating even more accountability in system.

Finally, common standards will make student expectations clear for teachers from year to year. The new standards will also enable more focused educator training and professional development. Effective, targeted teacher training is paramount, and common standards allow for teacher preparation programs and ongoing professional development to be focused on key objectives.

COMMON ASSESSMENTS DEVELOPMENT

Fewer, clearer, higher common core standards are only the first step in a longer reform process. The new standards lay the groundwork for States to collaborate on other key education reforms, including the development of next generation assessments. As States begin the standards adoption and implementation process, they

are also beginning the process of developing voluntary shared assessments, which will increase assessment quality, while also providing tremendous cost savings and other benefits. Aligned standards and assessments will allow States aligned teacher preparation and other supports designed to improve overall student achievement and close achievement gaps. Teachers from participating States will benefit from high quality instructional supports and materials that are aligned to the core standards.

CCSSO and NGA are providing support to two independent State assessment collaboratives and are working with several organizations to make sure that materials related to the common standards will be produced in an effective and open way to allow access to all teachers and schools. With common core standards and assessments, participating States can, as appropriate, continue collective reform efforts in nearly all facets of the education system.

THE FEDERAL ROLE

To preserve the project's integrity, it is imperative that the common standards initiative remains a State-led process. There are appropriate steps, however, the Federal Government can take to support State and local leadership. The revised Elementary and Secondary Education Act should reward State leadership and innovation, not just with funding for assessments, professional development, and other inputs, but also by codifying a new State-Federal partnership that promotes innovation and values State judgment on accountability. The current accountability system established under the No Child Left Behind Act will undercut movement toward high standards and must be updated to reflect the evolution of standards-based reform since the law was signed in 2002. By adopting the college and career-ready common core standards, States are voluntarily raising the bar for all students and the Federal Government should acknowledge their leadership by providing greater flexibility to help States ensure that all students meet these new higher expectations particularly as they transition their State accountability systems to the common core.

Thank you again for the opportunity to testify before the committee. I look forward to responding to any comments or questions you may have about this historic State-led effort.

The CHAIRMAN. Dr. Paine, thank you very much.

Now, Dr. Schmeiser, welcome. Please proceed.

STATEMENT OF CYNTHIA B. SCHMEISER, Ph.D., PRESIDENT, EDUCATION DIVISION, ACT NATIONAL OFFICE, IOWA CITY, IA

Ms. SCHMEISER. Good afternoon, Chairman Harkin, Ranking Member Enzi, and members of the committee.

Thank you for inviting me today to share some information about ACT's definition of college and career readiness, our role in the Common Core State Standards Initiative, and our approach to developing new generation assessments.

ACT, headquartered in your home State of Iowa, Chairman Harkin, is an independent, not-for-profit organization perhaps best known for the ACT, which is a widely used college admissions and placement assessment, but our scope and our mission has reached far beyond that to all levels of education and to workforce development, both nationally and internationally.

ACT's empirical research has defined college readiness as the acquisition of the knowledge and skills a student needs to enroll in and succeed in credit-bearing first-year courses at a post-secondary institution, such as a 2-year, a 4-year, or a trade school or technical school. Simply stated, readiness for college means not having to take remedial courses in college.

ACT research also shows that career readiness requires the same level of knowledge and skills in mathematics and in reading that college readiness does. Matter of fact, the majority of jobs that require at least a high school diploma, provide a living wage for a

family of four, are, in fact, projected to increase in number in the 21st century, and offer opportunities for career advancement require comparable levels of knowledge and skill for students entering workforce training programs as they do for the entering college student.

Compared to high school graduates who are not college ready, those who are ready to enter credit-bearing college courses are more likely to enroll in college, stay in college, earn higher grades, and graduate from college. Unfortunately, last year, of the 1.5 million students who graduated in 2009 and took the ACT, only 23 percent of those students were adequately prepared for post-secondary education in all four subject areas of English, math, reading, and science.

Because of our rich research base and our experience in college and career readiness, ACT was very pleased to play a major development role in the State-led common State standards initiative. The definition of college and career readiness within the Common Core State Standards Initiative is modeled on the approach pioneered by ACT.

And one of the most important distinguishing characteristics, I think, of this initiative is the fact that the standards were based on longitudinal research, research and evidence-based. It is an important distinction between the Common Core standards being research and evidence-based and what we have seen in State standards in this country. In addition to ACT's longitudinal research, the evidence was used from the work of high-performing countries, from high-performing States, as well as academic research.

Implementation of these standards provides a wonderful opportunity to better align and improve the educational and essential foundations of our system in this country around the goals of college and career readiness and represents a monumental step toward meeting our national goal of assuring education equity and excellence for all students. Along with the implementation, however, comes an important obligation to validate and to strengthen these standards periodically in an ongoing validation process.

Moving immediately from standards implementation to development assessment, I think, ignores some important steps, such as interpreting these standards into language that teachers can use and understand in the classroom and providing educators professional development in how to effectively teach these standards. Only then can these assessments, as part of an aligned, linked, and longitudinal system, be effective tools for students, teachers, administrators, and parents in monitoring student progress.

It is clear from ACT research that college and career readiness is a process. It is not a point in time. As such, no single assessment can effectively meet the needs of all. Therefore, ACT envisions States moving toward more coherent systems of assessment, comprised of multiple assessment measures and types.

Within such a system, each assessment would work with others to reveal a rich picture of student achievement and student growth. This will enable us to identify students who are on target, almost on target, and off target every grade, every year, allowing educators to tailor instruction to the needs of each student.

While the next generation assessments represent our highest expectations, we need to also remain sensitive to the pragmatic challenges faced by States, districts, and schools. It is important to strike an important balance between innovation and sustainability.

Therefore, based on our research, we would offer the following four recommendations. First, promote college and career readiness as a fundamental national goal and priority for all students.

Second, incentivize the implementation of college and career readiness standards by working with States to develop an accountability system that will meet their evolving needs.

Third, increase the capability of States, districts, and schools to use more effectively assessment data to monitor student progress, to intervene when students are falling behind, and differentiate instruction to advance college and career readiness for all students.

And finally, authorize additional resources for States implementing college and career readiness to develop coherent systems for assessment that include innovative measures, such as end-of-course, project-based learning and formative assessments.

We now have the opportunity to fill the promises that we have been making to our children for decades that when they graduate high school, they will be ready for college and work. ACT research has identified strategies that can help our Nation meet this goal, and we look forward to helping make college and career readiness a reality for each and every student through a reauthorized ESEA.

Thank you very much for this opportunity, and I look forward to answering any questions you might have.

[The prepared statement of Ms. Schmeiser follows:]

PREPARED STATEMENT OF CYNTHIA B. SCHMEISER, PH.D.

EXECUTIVE SUMMARY

ACT's empirical research defines *college* readiness as acquisition of the knowledge and skills a student needs to enroll and succeed in credit-bearing, first-year courses at a post-secondary institution, such as a 2- or 4-year college, trade school, or technical school. Simply stated, readiness for college means not needing to take remedial courses in college. ACT research also shows that *career* readiness requires the same level of knowledge and skills in mathematics and reading that college readiness does: the majority of the jobs that require at least a high school diploma, pay a living wage for a family of four, are projected to increase in number in the 21st century, and provide opportunities for career advancement require a level of knowledge and skills comparable to those expected of the first-year college student.

Because of our rich research base and expertise on college and career readiness, ACT was pleased to play a major development role in the State-led Common Core State Standards Initiative. The definition of college and career readiness within the Common Core State Standards Initiative is modeled on the approach pioneered by ACT. Implementation of those standards provides a remarkable opportunity to better align and improve the essential foundations of our Nation's education system around the goal of college and career readiness, and represents a monumental step toward meeting our National goal of ensuring educational equity and excellence for all students. Along with implementation comes the important obligation to validate and strengthen those standards periodically in an ongoing process. ACT will be working with States to help establish such a validation process.

The convergent timing of the development of the Common Core standards and the reauthorization of ESEA has spurred a productive national dialogue on how we can improve the purposes, design, and use of assessments in K–12 education. Moving immediately from standards implementation to the development of assessments ignores some important steps, such as interpreting those standards into language that teachers and leaders can understand and providing educators professional development on how to effectively teach the standards. Only then can assessments—as part of an aligned, linked, and longitudinal system—be effective tools for students, teachers, administrators, and parents in monitoring student progress.

No single assessment will effectively meet the needs of all. ACT envisions States moving toward more cohesive systems, comprised of multiple assessment measures and types. Within such a system, each assessment would work with the others to reveal a rich picture of student achievement and growth. This will enable us to identify students who are on target, nearly on target, or off target for college and career readiness in every grade, allowing educators to intervene with students who are falling behind.

Our Nation's efforts to strengthen standards and assessments will set high expectations for learning and provide educators with tools to monitor and accelerate student progress towards those expectations. This is a watershed moment in the history of education in our country. We at ACT look forward to helping make college and career readiness a reality for each and every student.

ABOUT ACT

Good afternoon, Chairman Harkin, Ranking Member Enzi, and members of the committee. Thank you for the opportunity to speak today about ACT's research on college and career readiness and its importance to the future of our Nation's students.

ACT, Inc. is an independent, not-for-profit organization that provides assessment, research, information, and program management services in education and workforce development. We are perhaps best known for the ACT test, the widely used college admission examination, but our scope and range go far beyond that one exam. Each year, we serve millions of individuals in middle schools, high schools, colleges, professional associations, businesses, and government agencies both nationally and internationally. Although designed to meet a wide array of needs, all ACT programs and services have one guiding purpose: helping people achieve education and workplace success.

For more than 50 years, ACT has collected and reported data on students' academic readiness for college by following millions of students into all types of post-secondary education to evaluate their success through college completion. ACT is the only organization with decades of data showing exactly what happens to high school graduates once they get to college or workforce training, based on how well they were prepared in middle school and high school.

While the attention paid to common college and career readiness standards and assessments is relatively recent, ACT has been implementing common standards and common assessments for well over 20 years. We have developed research-based standards that are linked to actual student success at the post-secondary level. As a result, the standards we have developed are generally fewer in number and more rigorous than what is typically found in many States' standards. In this model, our assessment data are comparable and transportable across State lines and have strong links to the post-secondary sector. As of the 2009–2010 school year, our College and Career Readiness System of vertically aligned assessments for 8th, 10th, and 11th or 12th grade students has been adopted statewide in 15 States and used at the district and school levels in all 50 States.

In this regard, our philosophy and approach are unique. Our assessments are grounded in research that tells us what knowledge and skills are *essential* in order for students to be ready for college and career. In my testimony I will share what we have learned and offer suggestions for reauthorization of the Elementary and Secondary Education Act (ESEA).

DEFINING COLLEGE AND CAREER READINESS

We are deeply encouraged by the national momentum to elevate the importance of college and career readiness within the policies and programs authorized under ESEA. While recognizing that the role of the Federal Government in education is limited, we do believe that ESEA can promote equity and excellence in communities nationwide through a unified goal of ensuring that every student leaves high school ready for college and career.

A first step toward realizing this goal is to come to agreement on what constitutes "college and career readiness." While there are many definitions of college and career readiness, the approach established through ACT research comes from empirical data.

ACT defines *college* readiness as acquisition of the knowledge and skills a student needs to enroll and succeed in credit-bearing, first-year courses at a post-secondary institution, such as a 2- or 4-year college, trade school, or technical school. Simply stated, readiness for college means not needing to take remedial courses in post-secondary education or training programs.

Unfortunately, there are far too many in this country who believe that the level of achievement needed for high school graduates who want to enter workforce training programs is far less than that needed for those students who plan to enter some form of post-secondary education. ACT research shows that *career* readiness requires the same level of foundational knowledge and skills in mathematics and reading that college readiness does. According to our research, the majority of the jobs that require at least a high school diploma, pay a living wage for a family of four, are projected to increase in number in the 21st century, and provide opportunities for career advancement require a level of knowledge and skills comparable to those expected of the first-year college student. The level of knowledge and skills students need when they graduate from high school is the same whether they plan to enter post-secondary education or a workforce training program for jobs that offer salaries above the poverty line.

What we have learned through our research is the critical importance that college and career readiness plays in college success. Compared to high school graduates who are not college- and career-ready, those who are ready to enter credit-bearing college courses are more likely to enroll in college, stay in college, earn good grades, and persist to a college degree. And in our latest research study soon to be released, we found that gaps in college success among racial/ethnic groups and by family income narrow significantly among students who are ready for college and career.

There is still much work to be done to ensure that all students graduate high school with this level of readiness. Of the 1.5 million high school graduates who took the ACT during academic year 2008–2009, 33 percent were not ready for college-level English, 47 percent were not ready for college social science, 58 percent were not ready for College Algebra, and 72 percent were not ready for college Biology. Overall, only 23 percent were ready to enter college-level courses without remediation in any of the four subject areas.

The remainder of my testimony will focus on the two issues at hand today: standards and assessments.

Allow me to first point out that the natural progression in building a cohesive, aligned educational system is not directly from standards to assessments, but rather from standards, to interpreting those standards into language that teachers and leaders can understand, to providing educators professional development on how to effectively teach the standards, to assessments that measure student progress linked to the standards, all followed by data monitoring and reporting to evaluate student progress and guide instruction. Therefore, I caution us to not make the assumption that standards and assessments—alone—are sufficient in and of themselves in ensuring college and career readiness for all students.

IMPLEMENTING THE COMMON CORE STATE STANDARDS

ACT has played a major role in the State-led Common Core State Standards Initiative, which seeks to articulate "fewer, clearer, and higher" K–12 education standards for voluntary adoption by States. The definition of college and career readiness within the Common Core State Standards Initiative is modeled on the approach pioneered by ACT. Endowed with extraordinary leadership from our Nation's governors and education chiefs, we believe that the Common Core initiative can be a catalyst for realizing the goal of preparing all students for college and career. I would like to address briefly some of the opportunities presented by this initiative.

In our view, the Common Core standards are of high quality, are easy to understand, and provide educators at the local level with the necessary flexibility to tailor instruction, curriculum, and professional development based on their own unique needs and contexts. The widespread enthusiasm for the draft common standards is a testament to the robust and open process that the initiative leaders established, and the hard work of many organizations and individuals from all over the Nation in developing, critiquing, and improving these standards.

One of the distinguishing characteristics of the Common Core State Standards Initiative is its insistence that evidence lead the way, rather than relying on subjective opinions about what students should be able to know and do when they leave high school. Not only did the initiative draw on ACT's longitudinal research on what knowledge and skills students need to succeed in post-secondary education and workforce training, but it sought additional evidence such as research from high-performing countries, standards from high-performing States, academic research on learning progressions, and other resources to support the inclusion of each and every standard within the Common Core framework.

Merely developing college and career readiness standards is not sufficient in and of itself. Along with development of standards comes the important obligation to validate and strengthen those standards periodically in an ongoing process. The

linkage between college and career readiness and success in post-secondary education and training, which has been a hallmark of ACT research, must now become a national priority. States have an obligation to ensure, through empirical validation—such as valuable feedback from post-secondary and workforce institutions to high schools about how well prepared their graduates were for college and career—that the level of readiness to which they are educating their students is continually being documented as sufficient preparation. ACT will be working with States to help establish such a validation process.

Further, our support of the Common Core initiative is predicated on the belief that this State-led movement provides a remarkable opportunity to better align and improve the essential foundations of our Nation's education system around this ambitious goal. We envision a future in which States, districts, and schools have fully aligned and integrated the core elements of their education infrastructure, including:

- expectations of what students need to learn and achieve through college and career readiness standards;
- instructional frameworks that broadly guide high-quality teaching and learning;
- rich and engaging classroom curricula and content;
- assessments aligned to college and career readiness standards and to what is taught in the classroom;
- systematic use of student data to improve teaching and learning;
- longitudinal data systems that enable the ongoing monitoring of student progress, allowing educators to identify students who are falling behind and accelerate them toward college and career readiness; and
- cohesive professional development programs for teachers and school leaders.

This opportunity to better align these elements—particularly in areas where there is a significant disconnect—would represent a monumental step toward meeting our national goal of ensuring educational equity and excellence for all students.

While ACT advocates for the better alignment of standards, curriculum, instruction, and assessment, we also fully realize that a one-size-fits-all model is unlikely to be successful given the remarkable diversity of our Nation's 15,000 school districts. Ultimately, the success of this initiative will rest with the educators and community members at the State and district levels who will be responsible for incorporating the standards into daily practice, making decisions about instruction and curriculum, and guiding each and every student toward college and career readiness.

Obviously, for many States and districts the transition to incorporating college and career readiness standards into daily practice will not happen overnight. We should recognize that many districts across the country will require additional capacity—both financial and human—to manage the transition to fewer, clearer, and higher standards.

For school, district, and State education leaders implementing the Common Core State Standards, there are several ways ESEA can provide critical assistance:

- promote college and career readiness as a fundamental national goal and priority for *all* students;
- support States, districts, and schools in developing monitoring systems that tell educators whether students are on target for college and career readiness at each grade level so that they can intervene when students fall behind academically; and
- incentivize the implementation of college and career readiness standards by working with States to develop an accountability system that will meet their evolving demands and allow for nuanced—not one-size-fits-all—evaluations of student achievement.

IMPROVING STATE ASSESSMENT SYSTEMS

The timing for the development of common standards and the reauthorization of ESEA has spurred a productive national dialogue on how we can improve the purposes, design, and use of assessments in K–12 education. ACT has used this opportunity to consult with State and local stakeholders to discuss what our own next-generation system should look like so that we can continue to be responsive to their current and future needs. I want to share some of what we have been learning from a wide variety of ACT stakeholders.

We know that no single assessment instrument is perfectly suited for meeting all of the purposes that teachers, education leaders, and policymakers have for assessment. When various assessment types are used in combination, they can provide a more comprehensive portrait of student and school progress than we have had in the past. We believe that it is possible to strike the appropriate balances among as-

sessment types to meet the multiple and varied needs of educators and policymakers while adhering to the highest professional standards.

We envision States moving toward more cohesive systems, comprised of multiple assessment measures and assessment types such as formative, interim, end-of-course, summative, and project-based assessments. While the widespread adoption of college and career readiness standards will help facilitate stronger alignment among the components of the assessment system, the assessments should also be designed from the start to be compatible with one another. Within such a system, each assessment would work with the others to reveal a richer picture of student achievement and growth, rather than operate in isolation. Such assessments enable us to identify students who are on target, nearly on target, or off target for college and career readiness, allowing educators to intervene with students who are falling behind.

The new generation of assessments should represent our highest aspirations while remaining sensitive to the pragmatic challenges faced by educators at the local and State levels: financial and human resources, access to necessary technology for computer-based testing, and educational practice. While the national dialogue on future assessment is focused on the promise of innovation, we recognize that even minor decisions about assessment design can have a significant impact on cost, complexity of administration, and scoring and reporting. In short, we need to strike an appropriate balance between innovation and sustainability.

What we have learned from State, district, and school leaders is informing ACT's development process as we move toward a next-generation assessment system. I hope that some of these lessons will be helpful to the committee in the reauthorization of ESEA:

1. College and career readiness is a process, not a single point in time. Growth and progress toward readiness must be monitored over a student's educational experience, starting in elementary school and through high school, so that timely instructional decisions and interventions can be made.

2. Assessments need to be part of a system that is aligned, linked, and longitudinal in nature if it is to be an effective tool for students, teachers, administrators, and parents in monitoring student progress. We must be exceptionally clear in defining the purposes, uses, and limits of effective assessment.

3. State assessment systems should include not only measures of academic achievement and growth, but also measures of those academic behaviors that influence readiness and educational and career planning.

4. The unique needs of English Language Learners and students with disabilities should be incorporated from the start of the assessment design process and with the deep consultation of stakeholders and experts.

5. Assessment formats should be varied according to the type of achievement that needs to be measured. These multiple measures can be used to offer more comprehensive evaluations of student achievement, from multiple-choice and constructed-response assessments to project-based learning.

6 Assessment should be offered through multiple platforms. While computer-based testing is highly applicable to formative assessments that can be conducted on an on-demand basis, paper-and-pencil testing may be a reality for States and districts with less technological capacity. Until computer access for such large groups of students is more available in schools, we need to use both platforms flexibly and wisely.

7. Ongoing, real-time, interactive reporting and access to data by multiple stakeholders—especially teachers—is essential if stakeholders are to get the most out of assessment results.

Given our experience at implementing high-quality assessments tied to college and career readiness standards, ACT offers the following recommendations for the State assessment component of ESEA reauthorization:

- continue to improve summative State assessments for the purposes of student monitoring and accountability measured against the standards;
- authorize additional resources for States implementing college and career readiness standards to develop coherent systems of assessment that include innovative measures such as end-of-course, project-based, and formative assessments; and
- increase the capability of States, districts, and schools to more effectively use assessment data to monitor student progress, intervene when students are falling behind, and differentiate instruction to advance college and career readiness for *all* students.

Taken together, our Nation's efforts to strengthen standards and assessments will be a critically important accomplishment, but are merely two essential pieces of the puzzle. Improvements to standards and assessments will not in and of themselves result in dramatic improvements in student outcomes. Rather, they set high expec-

tations for learning and provide educators with tools to monitor student progress towards those expectations. What we have learned from high-performing countries and high-performing districts domestically is that, in order to succeed at improving the college and career readiness of our students, we must develop an aligned and coherent system of standards, curriculum and instruction, assessment, professional development, and student support programs, with all of these components contributing to an authentic process of continuous improvement in all phases of daily educational practice.

To say that we are experiencing a watershed moment in the history of education in our country is an understatement. We are poised to make incredible progress in advancing the preparation of our Nation's students for college and career. We have an opportunity to fulfill the promises we have been making to our children for decades—that when they graduate from high school they will be ready for college and work. ACT's research has identified strategies that can help our Nation meet this goal. There is still much to be done, and a reauthorized ESEA can help accomplish it. We look forward to helping make college and career readiness a reality for each and every student.

I look forward to answering any questions you may have.

REFERENCES

Focusing on the Essentials for College and Career Readiness—Policy Implications of the ACT National Curriculum Survey Results 2009 (*http://www.act.org/research/policymakers/pdf/NCS*__PolicySummary2009.pdf).

College Readiness—ACT Research on College and Career Readiness: A Summary of Findings (*http://www.act.org/research/policymakers/pdf/NCS*__PolicySummary2009.pdf).

The CHAIRMAN. Thank you, Dr. Schmeiser, very much.

Now we turn to Dr. Phillips. Dr. Gary Phillips.

STATEMENT OF GARY PHILLIPS, Ph.D., VICE PRESIDENT, AMERICAN INSTITUTES FOR RESEARCH, WASHINGTON, DC

Mr. PHILLIPS. Thank you, Chairman Harkin, Ranking Member Enzi, and members of the committee, for the invitation to be here today.

My name is Gary Phillips. I am a vice president and chief scientist at the American Institutes for Research.

I would like to make two central points about No Child Left Behind and the reauthorization of the ESEA. First, No Child Left Behind has a large loophole that has misled the public, and I encourage Congress to close this loophole in the reauthorization of the ESEA. Second, Congress should encourage States to abandon their outmoded 20th century tests for a new generation of technology-based tests that more accurately measure growth, that are less burdensome, that are faster and cheaper.

To my first point, the most significant thing wrong with No Child Left Behind is a lack of transparency. The consequences of failing to meet adequate yearly progress had the unintended consequence of encouraging States to lower, rather than raise, their own standards. The law inadvertently encouraged the States to dumb down their performance standards to get high rates of proficiency.

The fact that States dumb down their performance standards can be seen in the figures, Figures 1 and 2, in my full statement that I provided to you. The percent proficient in these graphs represent what was reported by No Child Left Behind in mathematics in 2007. Using Grade 8 as an example, according to No Child Left Behind, Tennessee is the highest-achieving State in the Nation, while Massachusetts is one of the lowest.

There is something wrong with this picture. According to NAEP, exactly the opposite is true. Massachusetts is the highest-achieving

State in the Nation, with Tennessee being one of the lowest. I say this with due respect to Senator Lamar Alexander.

If we look deeper into the State standards, we begin to explain this contradiction. The grades imposed on the charts in the figures I provided to you are from an upcoming report from AIR titled "The Expectations Gap" that internationally benchmarked the State standards and internationally benchmarked those standards to the Trends in International Mathematics and Science Study, TIMSS. In other words, it benchmarked the standards internationally.

Returning to Grade 8, in my full statement, we see that many States obtain high levels of proficiency by lowering their standards. The States with the highest levels of proficiency require only a D, which is comparable in difficulty to the lowest level of mathematics on TIMSS.

In fact, the correlation between the percent proficient and the level of the standard is negative 0.8. This means as States lower their standards, they raise their level of reported proficiency.

The difference in the State standard is not just a minor accounting irregularity. It has real equity consequences for students' opportunity to learn. If my child attends school in a State where almost everyone is proficient, what leverage do I have as a parent to ask the State to provide a more challenging education?

How big is this expectation gap, the difference between the highest and lowest standards across the State? This gap is more than twice the size of the national black-white achievement gap. The Nation will never be able to close the achievement gap until it reduces this bigger gap and all States adopt higher standards.

This helps explain why we do so poorly on international comparisons. Many States think they are doing well and feel no urgency to improve because almost all their students are proficient. They have no idea how they stack up when compared to peers outside of their own borders.

And now to my second point. The outmoded pencil and paper tests used in most States are costly and time consuming. States claim they teach 21st century skills, but they measure learning with 20th century tests. The only way States will modernize and take advantage of high-speed technology is with Federal funding.

Furthermore, the outmoded testing paradigm provides poor measurement for a large portion of students in the population. These tests are too easy for the highest achieving students, and they are too hard for the lowest achieving students, especially students with disabilities and English language learners.

The $350 million from the Race to the Top assessment fund and the reauthorization of the ESEA could provide an unprecedented opportunity for States to upgrade their testing capacity. I would recommend that the ESEA encourage the future consortia of States to use computer-adaptive testing as their standard modus operandi.

These types of tests are already in partial use in many States. However, in three States—Delaware, Hawaii, and Oregon—the entire State testing program is already computer adaptive. Since AIR is the vendor in these three States, I can speak with some authority on how these tests operate.

In all three of these States, the tests consist of multiple choice and challenging constructed response items that are both administered and scored by the computer. There are no printing costs, no scoring costs. In fact, the long-run total cost of the system is half that of paper and pencil tests.

In each of these three States, the test is developed based on universal design principles, and the test content is the same for each student that is tested. The technology platform provides three opportunities to take the test each year. In addition, teachers can develop their own formative assessment, and interim assessments are also provided, all computer adaptive and all on the same scale as the summative test. The results are available for each student within 15 seconds.

In conclusion, I would like to thank you for the opportunity to give you my views on the next generation of State assessments. Setting internationally competitive education standards is a critical national priority.

Students tomorrow will not be competing with the best students in their school. They will be competing with the best students in the world. In order to get States to establish high standards, you must close the expectations loophole in No Child Left Behind and reward States that set high internationally benchmarked standards.

States also need Federal funding in order to embrace the next generation of technology-driven assessments. The technology for better, faster, and cheaper testing is already here. National leadership is needed to move the States in this direction.

Thank you very much.

[The prepared statement of Mr. Phillips follows:]

PREPARED STATEMENT OF GARY W. PHILLIPS, PH.D.

SUMMARY

I would like to make two central points about No Child Left Behind and the reauthorization of the ESEA. (1) No Child Left Behind has a large loop hole that has misled the public and I encourage Congress to close this loop hole in the reauthorization of the ESEA. (2) I will propose that Congress encourage States to abandon their outmoded 20th century paper/pencil tests for a new generation of 21st century technology-based tests that are more accurate, less burdensome, faster, and cheaper.

The most significant thing wrong with NCLB is a lack of transparency. The severe consequences of failing to meet AYP had the unintended consequence of encouraging States to lower, rather than raise, their own standards. The law inadvertently encouraged the States to dumb down their performance standards to get high rates of proficiency. The fact that States dumb down their performance standards can be seen in Figures 1 and 2 in this document. The "percent proficient" in these tables represent what was reported by NCLB in Grades 4 and 8 in mathematics in 2007. Using Grade 8 as an example, we see that Tennessee is the highest achieving State in the Nation while Massachusetts is one of the lowest. However, if we look deeper into State performance standards, we see a different story. The grades imposed on the chart are from an upcoming AIR report titled "The Expectation Gap" that internationally benchmarked State proficient standards to the Trends in International Mathematics and Science Study (TIMSS). Returning to Figures 1 and 2, we see that many States obtain high levels of proficiency by lowering their standards. The States with the highest levels of proficiency require only a *D*, which is comparable to the lowest level of mathematics knowledge and skills on TIMSS. In fact, the correlation between the percent proficient reported by the State and the difficulty of their standards is $-.81$. The gap in expectations in the State performance standards is not just a minor accounting irregularity. It has real equity consequences for a student's opportunity to learn. If my child attends school in a State where almost everyone is proficient, what leverage do I have as a parent to ask the State to provide

a more challenging education? The expectations gap has major educational consequences. This *expectation gap* is so large that it is more than twice the size of the national black-white *achievement gap*. The Nation will never be able to close the achievement gap until it closes the bigger problem of the expectations gap. This helps explain why the United States does poorly on international comparisons. Many States think they are doing well and feel no urgency to improve because almost all their students are proficient. They have no idea how they stack up when compared to peers outside their own Lake Woebegone. This also helps explain why almost 40 percent of students entering college need remedial courses. They thought they were college ready because they passed their high school graduation test—but they were not.

The outmoded paper/pencil tests used in most States are costly and time consuming. States claim they teach 21st century skills but they measure learning with 20th century tests. The only way State testing will move into the 21st century and take advantage of high-speed modern technology is with Federal funding. Furthermore, the current model of one-size-fits-all, paper/pencil test provides poor measurement for much of the student population. The tests are too easy for high-achieving students and too hard for low-achieving students, students with disabilities, and English language learners. The $350 million from the Race to the Top Assessment Program and the reauthorization of the ESEA could provide an unprecedented opportunity for States to upgrade their testing capacity. I would recommend that the ESEA encourage the consortia of States to use *Computer-Adaptive Testing* as their standard modus operandi. *Computer-adaptive tests* are already in partial use in many States. However, in three States—Delaware, Hawaii, and Oregon—the entire State testing program is already computer-adaptive. In all three of these States, the test consists of multiple-choice items and challenging constructed-response items that are both administered and scored by computer (no printing cost and no scoring cost). The total cost of the computer-adaptive test is half that of a paper/pencil test. In each of these three States, the computer-adaptive test is developed based on universal design principles, and each test administered to a student covers all of the content standards. The technology platform provides three opportunities to take the summative test each year (used for accountability and Federal reporting). In addition, the computer-adaptive test administers teacher-developed formative assessments and interim assessments all on the same scale as the summative test. The results are available for each student within 15 seconds.

Thank you Chairman Harkin and members of the committee for the invitation to be here today. My name is Gary W. Phillips, and I am a Vice President and Chief Scientist at the American Institutes for Research (AIR). AIR is a 65-year-old, not-for-profit, nonpartisan organization whose mission is to conduct behavioral and social science research to improve people's lives and well-being, with a special emphasis on the disadvantaged. Previously, I was the Acting Commissioner at the National Center for Education Statistics (NCES). My career has been devoted to providing policymakers with better data to help them improve American education.

Today I would like to make two central points about No Child Left Behind and the reauthorization of the ESEA.

1. I will demonstrate that No Child Left Behind (NCLB) has a large loop hole that has misled the public and I encourage Congress to close this loop hole in the reauthorization of the ESEA. Other people will be providing you testimony on whether this legislative act will improve education. I will focus on whether this legislative act provides enough information to know if education has been improved.

2. I will propose that Congress encourage States to abandon their out-moded 20th century paper/pencil-based testing paradigm for a new generation of 21st century technology-based tests that are more accurate, less burdensome, faster, and cheaper.

WHAT IS WRONG WITH NO CHILD LEFT BEHIND?

The most significant thing wrong with NCLB is a lack of transparency. Contributing to this lack of transparency is the fact that the NCLB results represent State efforts to reach unattainable national goals. For the last quarter century, education reform professionals have known that our underachieving educational system has put our Nation at risk (*A Nation at Risk: The Imperative for Education Reform,* April 1983). National policymakers have responded to this crisis with slogans and unattainable utopian goals, such as "*being the first in the world in mathematics and science achievement by 2000*" (1990 National Education Goals Panel); or "*all students will be proficient in reading and mathematics by 2014*" (No Child Left Behind Act of 2001); or "by 2020 . . . *ensure that every student graduates from high school*

well prepared for college and a career" (A Blueprint for Reform: The Reauthorization of the Elementary and Secondary Education Act, 2010). A national goal should be high but reachable. A good example of a challenging but achievable national goal is the *Proficient* standard used by the National Assessment of Educational progress (NAEP) and the National Assessment Governing Board (NAGB). The *Proficient* standard is challenging but achievable by most (although not all) students. The new ESEA should contain career and college-ready national goals that are internationally competitive but not so high that they are unattainable by States and schools.

The greatest contributor to the lack of transparency in NCLB, however, is the misleading data used by policymakers to monitor progress toward the goals (referred to as Adequate Yearly Progress). Both the Federal Government and the States have an unfortunate history of presenting flawed State testing data to the public.

From 1984 to 1989, the U.S. Department of Education compared State performance using the Wall Chart that showed average State aggregates of SAT and ACT scores. The Wall Chart was used even though it was widely criticized because it measured only the self-selected college-bound population. The larger the percentage of the population taking the SAT or ACT tests, the lower the State's ranking on the Wall Chart. The States with the least number of students heading for college tended to have the highest ranking. In fact, the 1986 correlation between the SAT and the proportion of college-bound students was -0.86 (College Board, 1986). The fact that it was a misleading indicator due to self-selection did not deter the department from using the system for 6 years under two Secretaries of Education, Terrell H. Bell and William J. Bennett.

In 1987, a West Virginia physician produced a report in which he stated that he had found that on norm-referenced tests, all 50 States were claiming they were above the national average (Cannell, 1987). This so-called Lake Woebegone report sparked much interest in Washington because it was hoped that norm-referenced tests might overcome some of the problems of the SAT and ACT in the Wall Chart as indicators of State-by-State performance. Although this was a black eye for educators, the practice continues today. States are still asked to explain how they can be above the national average on their norm-referenced test when they are below the national average on the National Assessment of Educational progress (NAEP).

The biggest flaw in State testing data, however, is in use today in all States, sanctioned and encouraged by the No Child Left Behind Act of 2001. NCLB provides a new type of Wall Chart where again State aggregates are not comparable and are misleading. The most significant thing wrong with NCLB is a lack of transparency. The severe consequences of failing to meet AYP had the unintended consequence of encouraging States to lower, rather than raise, their own standards. The law inadvertently encouraged the States to dumb down their performance standards to get high rates of proficiency. The fact that States dumb down their performance standards can be seen in Figures 1 and 2 in this document. The "percent proficient" in these tables represent what was reported by NCLB in Grades 4 and 8 in mathematics in 2007. In my remaining remarks I will use Grade 8 to illustrate my points. In Grade 8 we see that Tennessee is the highest achieving State in the Nation while Massachusetts is one of the lowest. If parents were looking to raise a family in a State with an excellent track record of success based on NCLB data, they should move their family to Tennessee. However, there is something wrong with this picture. We know that NAEP reports exactly the opposite with Massachusetts the highest achieving State and Tennessee being one of the lowest achieving States.

However, if we look deeper into State performance standards, we can begin to explain this contradiction. The grades imposed on the chart are from an upcoming AIR report titled "The Expectation Gap" that internationally benchmarked State proficient standards to the Trends in International Mathematics and Science Study (TIMSS) (Phillips, 2010). The report then expressed the international benchmarks as international grades. To do this I statistically linked the test in each State to the Trends in International Mathematics and Science Study (TIMSS) and expressed the State standards as international grades on a comparable scale. (A = Advanced, B = High, C = Intermediate, D = Low). This gives policymakers an international benchmarked common metric by which to compare State performance standards. Returning to Grade 8 we see that many States obtain high levels of proficiency by lowering their standards. The States with the highest levels of proficiency require only a D, which is comparable to being below the *Basic* standard on NAEP and the lowest level of mathematics knowledge and skills on TIMSS. On the other hand, the States with the lowest levels of proficiency require the highest standards (where a B is comparable to the *Proficient* standard on NAEP and equal to the *High* level on TIMSS). In fact, the correlation between the percent proficient reported by the State under NCLB and the difficulty of their standards is $-.81$.

The gap in expectations in the State performance standards is not just a minor accounting irregularity. It has real equity consequences for a student's opportunity to learn. If my child attends school in a State where almost everyone is proficient, what leverage do I have as a parent to ask the State to provide a more challenging education? The gap in expectations has major educational consequences. The difference between the standards in Massachusetts and the standards of the States with the lowest standards is about two standard deviations. This gap in expectations is so large that I would like to take a minute to impress on you just how large it is.

1. This *expectation gap* is so large that it is more than twice the size of the national black-*white achievement gap*. The Nation will never be able to close the achievement gap until it closes the bigger problem of the expectations gap.

2. The gap in expectations represents two-to-three grade-level differences between what the States are expecting their students to know and be able to do. What the low-standard States are expecting in middle school is comparable in difficulty to what Massachusetts expected back in elementary school.

3. The Massachusetts proficient standard is at the 54th percentile. If Massachusetts used the Tennessee proficient standard in Massachusetts it would be at the 4th percentile.

This helps explain why the United States does poorly on international comparisons. Many States think they are doing well and feel no urgency to improve because almost all their students are proficient. They have no idea how they stack up when compared to peers outside their own Lake Woebegone. This also helps explain why almost 40 percent of students entering college need remedial courses. They thought they were college-ready because they passed their high school graduation test—but they were not.

We should note that not all States are achieving high rates of proficiency by lowering their standards. For example, Hawaii is a small and relatively poor State that has made the right policy decision that is in the best interest of its children by requiring high standards (just under those in Massachusetts), although student performance is relatively low. Even though the State has been internally criticized for having too high standards, the State leadership has maintained the high standards and the student's performance in Hawaii have gradually improved (as indicated by their NAEP scores) over the years.

Figure 1: How High Were the State Proficient Standards in 2007 in Grade 4 Mathematics?

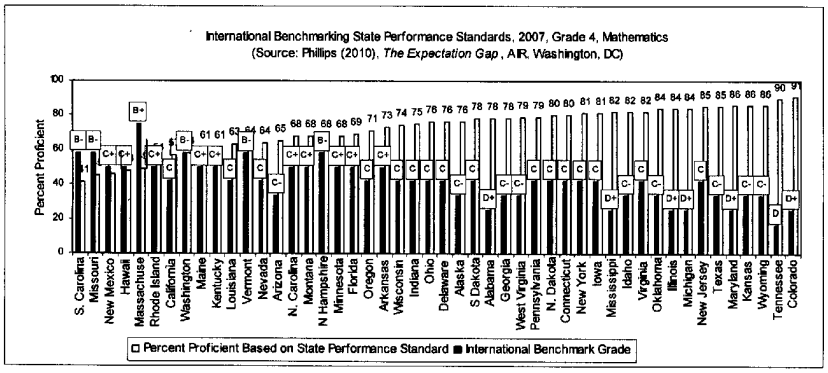
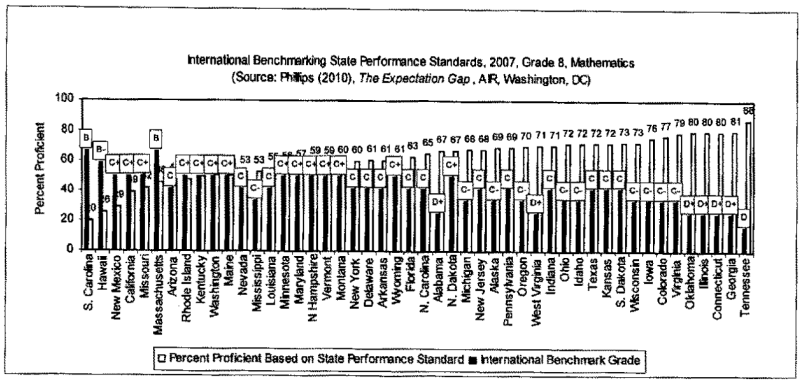
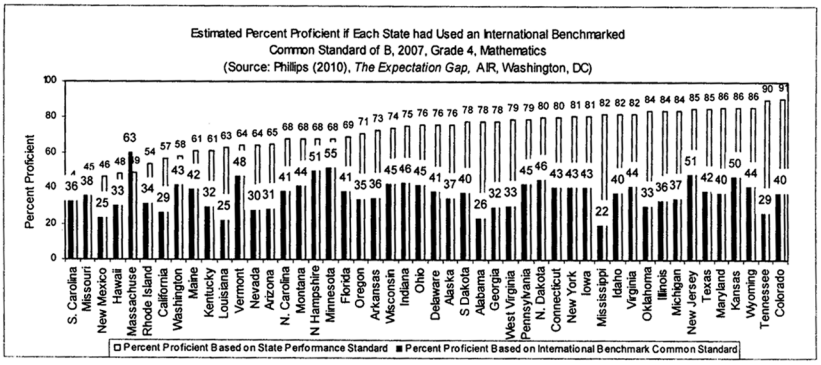


Figure 2: How High Were the State Proficient Standards in 2007 in Grade 8 Mathematics?



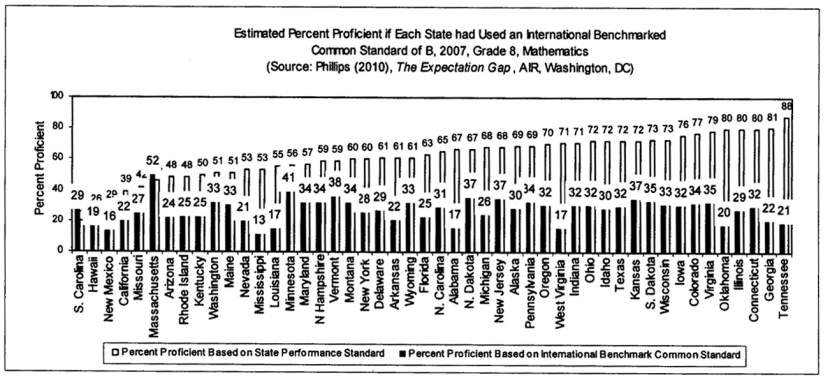
How would the 2007 State results reported to NCLB have looked had all the States used a *common performance standard* that was comparable in difficulty to the High International Benchmark on TIMSS? Had this been done, then all of the States would have reported their percent proficient based on performance standards of comparable difficulty using a level playing field. Figure 4 gives an example of what this might have looked like for Grade 8 mathematics—a dramatically different picture of State performance. We see that when all the States use an internationally competitive *common performance standard*, the performance in Tennessee drops from 88 percent to 21 percent. Now Massachusetts is the highest achieving State. If the parents mentioned above were using the information shown in Figure 4 to choose a State in which to live, where their children would attend schools with the highest educational expectations and achievement, they might choose Massachusetts.

Figure 3: How Many Students Would Have Been Proficient if Each State Had Used a Level Playing Field in 2007 Grade 4 in Mathematics?



56288-2.eps 56288-3.eps

Figure 4: How Many Students Would Have Been Proficient if Each State Had Used a Level Playing Field in 2007 Grade 8 in Mathematics?



THE NEED FOR A NEW GENERATION OF TECHNOLOGY-BASED STATE TESTING

NCLB requires that States develop their own tests but does not provide funding for doing so. Therefore, States suffering budget cutbacks have no incentive to try new and better approaches to testing. The outdated pencil/paper tests used in most States require costly and time-consuming administration, followed by costly and time-consuming scoring, followed by costly and time-consuming reporting. With spring testing, getting test results back to teachers and parents before the summer recess is nearly impossible. States like to claim they teach 21st century skills but they measure learning with 20th century tests. The only way State testing will move into the 21st century and take advantage of high-speed modern technology is with Federal funding. Furthermore, the current model of one-size-fits-all, paper/pencil test provides poor measurement for much of the student population. The tests are too easy for high-achieving students and too hard for low-achieving students, students with disabilities, and English language learners.

The $350 million from the Race to the Top Assessment Program and the reauthorization of the ESEA could provide an unprecedented opportunity for States to upgrade their testing capacity. In the near future, many States are likely to function as consortia and adopt the Common Core Standards developed by the Council of Chief State School Officers (CCSSO) and the National Governors Association (NGA). I would recommend that the ESEA encourage the consortia of States to use *Computer-Adaptive Testing* as their standard modus operandi.

Computer-adaptive tests are already in partial use in many States. However, in three States—Delaware, Hawaii, and Oregon—the entire State testing program are already computer-adaptive. Since AIR is the vendor for these three States I can speak with some authority on how their computer-adaptive tests operate. In all three of these States, the test consists of multiple-choice items and challenging constructed-response items that are both administered and scored by computer (no booklet printing cost and no scoring cost). The total cost of the computer-adaptive test is half that of a paper/pencil test. In each of these three States, the computer-adaptive test is developed based on universal design principles, and each test administered to a student covers all of the content standards. The technology platform provides three opportunities to take the summative test each year (used for accountability and Federal reporting). In addition, the computer-adaptive test administers formative assessments (developed and used by teachers for diagnostic purposes) and interim assessments (used by teachers to get an early fix on how much students are progressing during the year) all on the same scale as the summative test. The results are available for each student within 15 seconds. Not only are these assessments faster and cheaper, but computer-adaptive testing yields more accurate measurement for high- and low-achieving students and better measurement for students with disabilities and English language learners.

56288-4.eps

WHAT SHOULD BE INCLUDED IN THE REAUTHORIZATION OF ESEA?

Common content standards and common performance standards should be included in the reauthorization of ESEA. The CCSSO and the NGA are currently developing common content standards. Content standards represent the scope and sequence of content that should be taught in the schools. This is an important first step in creating transparency and accountability in ESEA. However, this needs to be followed by an equally important second step—establishing common performance standards. Performance standards represent how much, of what is taught, students are expected to learn. Because every student cannot learn everything that is taught in every grade and every subject, educators need a realistic performance goal. This performance standard needs to be common to all the States (or consortium of States) so that all the States have a level playing field. Each State does not get to set its own bar. The United States cannot be internationally competitive in our educational achievement if States are going in 50 different directions (different content standards) and have 50 different expectations of what their students should learn (different performance standards).

Computer-adaptive testing and the use of the best available modern technology should be included in the reauthorization of the ESEA. The reauthorized ESEA should encourage and fund States to use modern technology to administer, score, and report results. The best of all options is computer-adaptive testing that provides a more reliable measurement of student achievement involving less time, fewer items, and less cost. Computer-adaptive testing also provides better measurement for both high-achieving students and low-achieving students such as students with disabilities and English language learners.

In conclusion, I would like to thank you for the opportunity to give you my views on the next generation of State assessments. Setting internationally competitive education standards is a critical national priority. Students tomorrow will not be competing with the best students in their school. They will be competing with the best students in the world. In order to get States to establish high standards you must close the expectations loop hole in NCLB and reward States that set high internationally benchmarked standards. States also need Federal funding in order to embrace the next generation of technology-driven assessments. The technology for better, faster and cheaper testing already exists. National leadership is needed to move the States in this direction.

REFERENCES

Bandeira de Mello, V., Blankenship, C., and McLaughlin, D. (2009). *Mapping State Proficiency Standards onto NAEP Scales: 2005–2007* (NCES 2010–456). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Cannell, J.J., (1987), Nationally Normed Elementary Achievement Testing in America's Public Schools: How All 50 States Are Above the National Average, Friends for Education, Daniels, WV.

College Board, (1986), *Press statement for release of 1986 SAT scores*. New York: The College Board.

Phillips, G.W., (2010), The Expectation Gap: Internationally Benchmarking State Performance Standards, American Institutes for Research, Washington, DC.

The CHAIRMAN. Well, Dr. Phillips, thank you very much for that thought-provoking statement.

Next, we turn to Dr. Rivera. Dr. Rivera, welcome. Please proceed.

STATEMENT OF CHARLENE RIVERA, Ed.D., EXECUTIVE DIRECTOR, GEORGE WASHINGTON UNIVERSITY CENTER FOR EQUITY AND EXCELLENCE IN EDUCATION, ALEXANDRIA, VA

Ms. RIVERA. Thank you.

Good afternoon, Chairman Harkin, Ranking Member Enzi, and members of the HELP Committee. I am very pleased to be here today to have this opportunity to speak to you about this issue.

I am Charlene Rivera, research professor at the George Washington University and executive director of the Center for Equity and Excellence in Education at the university.

Many years ago, I was a bilingual teacher in the Boston Public Schools and now have the opportunity to conduct research that relates to English language learners.

I am pleased to offer my perspective on how the Common Core standards and assessments should address and measure academic outcomes for English language learners. I also would like to address the challenges faced in developing assessments that can support teaching and learning for these students.

Initially, however, it is very important for all of us to really know who these English language learners are. And while constructing a coherent system of standards, instruction, and assessment that can address all these students, including English language learners, it is important to take into account the need for the Common Core standards and new assessment system or systems to recognize and address the linguistic needs of these students.

To recognize that English language learners need instruction in academic language to acquire subject matter proficiency—and I am defining academic language as the language that is used in school to help students acquire and use knowledge—to acknowledge that English language proficiency standards and assessments are distinct from English language arts standards and assessments, and finally, to recognize that implementation is key to success of the new system.

English language learners are not a homogenous group, and attention to their different characteristics is essential to helping them succeed and be college ready. These students differ in their level of English language proficiency, ethnic background, socioeconomic status, quality of prior schooling, and literacy level in their first language.

Many English language learners are economically and educationally disadvantaged and attend high-poverty schools. These schools often lack the educational resources and personnel knowledgeable about how to teach these students. Effectively educating English language learners requires adjusting or differentiating instructional approaches, content instruction, and assessment.

The Common Core standards and new assessment system must address the linguistic needs of these students. However, because the English language arts standards are developed with native English-speaking students in mind, it is important to consider the role and use of English language proficiency standards and assessments.

It will be important to articulate the relationship and to clearly delineate expectations for when instruction in English language arts versus English language proficiency is appropriate for English language learners. This specification should be established in every State or consortium of States by a working committee of English language learner and English language arts experts.

This group should use data to determine at what point along the continuum of learning English, English language learners at low levels of English language proficiency should be held accountable for English language arts standards. For these students, it is seriously worth considering substituting the English language proficiency reading and writing standards and assessments as meas-

ures of reading and writing achievement, at least for a short period of time.

With regard to the Common Core mathematic standards, it is important to consider whether these students need to be addressed only in English—or whether these standards need to be addressed only in English or if they can also be addressed in students' native languages.

Successful implementation of the new system requires changes to teacher preparation and in-service professional development programs. These programs must build the capacity of content and English as a second language for teachers to differentiate instruction and classroom assessment and, in addition, to teach the academic language required for English language learners to be successful in academic content.

English language learner experts must be involved at every level of design and implementations. States should consider the needs of English language learners in the new standards and assessment system. State policies must address how these students are identified and address procedures for including and accommodating students in summative benchmark and classroom assessments. Most importantly, the new assessment system must be valid and reliable for all students, including English language learners.

At the Federal level, the Department of Education needs to improve the review and monitoring of the standards and consortium assessment systems. It is crucial that the review processes explicitly address English language learners and that the reviewers have the necessary expertise and knowledge to evaluate the adequacy of the assessment system for these students.

In conclusion, the design of assessment and accountability systems and their implementation must address the linguistic diversity and other characteristics of English language learners. To be successful, the system must ensure that the standards and assessment processes address academic language as well as English language proficiency. It must be recognized that academic language is a barrier for English language learners and needs to be taught explicitly.

I have great expectations for the ESEA reauthorization and look forward to an interconnected system of standards, instruction, and assessment that works for all students, including English language learners.

Thank you for the opportunity to speak today, and I will be happy to answer any questions.

[The prepared statement of Ms. Rivera follows:]

PREPARED STATEMENT OF CHARLENE RIVERA, ED.D.

SUMMARY

It is imperative that the needs of English language learners (ELLs) are addressed in the reauthorization of ESEA. ELLs are not a homogenous group and attention to their different characteristics is essential to meaningfully instructing and assessing them.

Although English language proficiency (ELP) and English language arts (ELA) are related and even list the same skills (listening, reading, and writing), presumptions about students' background and basic competencies in English differ. For ELLs at low levels of ELP it is worth considering substituting the ELP reading and writing standards and assessments as measures of their reading and writing achievement.

A crucial factor for ELLs to meet standards is being able to understand and use the academic language or academic English of different disciplines. While a mastery of academic language is demanding for all students, it can be especially difficult for students who already struggle with other linguistic challenges, such as ELLs and former ELLs. In a reauthorized ESEA, resources should be allotted to States to work toward the development of a broad national framework that captures the many dimensions of academic English.

States should consider the needs of ELLs in the new standards and assessment system. Policies must address how ELLs are defined, and address procedures for including and accommodating them in summative, benchmark, and classroom assessments. There is great need to clearly distinguish the linguistic needs of ELLs from cognitive, processing, or physical needs of students with disabilities. The delineation of policy at the State and consortium levels is important and should guide practice for the new assessment system which must be valid and reliable for all students including ELLs.

At the Federal level, ED must improve the review and monitoring of the standards and assessment systems. It is crucial that the review processes explicitly address ELLs and that the reviewers have the appropriate expertise and knowledge.

In conclusion, the design of assessment and accountability systems and their implementation must consider the linguistic diversity and other characteristics of ELLs. To be successful, the system must ensure that the standards and assessment processes address academic language as well as English language proficiency. Teacher preparation and in-service professional development programs must build the capacity of content and ESL teachers to differentiate instruction and assessment for ELLs, as well as teach ELLs the academic language required to successfully access the academic content. ELL experts must be involved at every level of design and implementation. I have great expectations for the ESEA reauthorization and look forward to an interconnected system that works for English language learners.

Chairman Harkin, Ranking Member Enzi, and members of the HELP committee, thank you for inviting me to testify today and offer my perspective on how the common core standards and assessments should address and measure academic outcomes for English language learners, or ELLs. I also will address the special challenges faced in developing assessments which provide information that can support teaching and learning for ELLs.

Initially, however, it is important that we clearly define and have a shared understanding of ELLs. Also while constructing a coherent system of standards, instruction, and assessment that can address all students including ELLs, it is important to take into account:

- the need for the Common Core Standards and new assessment system(s) to recognize and address the linguistic needs of ELLs;
- that ELLs need academic language to acquire subject matter proficiency; and
- that English language proficiency (ELP) standards and assessments are distinct from English language arts (ELA) standards and assessments.

U.S. schools serve over 5 million ELLs. These learners are scattered across the United States and are highly mobile. About 10 to 12 percent of students in public schools are ELLs. While the number of ELLs continues to increase in Northeast and Western States that traditionally have had large numbers of ELLs, more recently, the Southeast and Midwest have seen dramatic increases. The impact of these demographic changes on schools makes it imperative that the needs of the ELL population are addressed in the Blueprint and supporting proposals guiding the reauthorization of ESEA.

ELLs are not a homogenous group and attention to their different characteristics is essential to meaningfully instructing and assessing them. One important example is the level of English language proficiency, but ELLs also differ in ethnic background, socioeconomic status, quality of prior schooling, and first or native language, including literacy in their first language. Many ELLs are economically and educationally disadvantaged and attend high-poverty schools. All too often the schools ELLs attend lack the educational resources and personnel knowledgeable about how to teach them the academic English or academic language needed to acquire the content knowledge and skills needed to reach high academic standards, graduate from high school, and be college ready.

As Short and Fitzsimmons (2007) argue, ELLs must "perform *double the work* of native English speakers in the country's middle and high schools" (p. 1) because they are studying content area subjects through a language in which they are not

yet fully proficient. In order to understand and apply academic concepts, students must be able to interpret and produce complex oral and written language.

Effectively educating these students requires adjusting or differentiating instructional approaches, content instruction, and assessment in ways that take into consideration their differences. However, practices for identifying who is an ELL are not systematic across or sometimes even within States. Therefore, one of the basic issues to address in a reauthorized ESEA is clearly defining the ELL subgroup by requiring all schools and districts within a State to apply comparable screening, entry, and exit criteria.

As recommended by the Working Group on ELL Policy, a reauthorized ESEA should require States to establish stable ELL subgroup membership for accountability purposes (see Working Group on ELL Policy Recommendations at *http://ellpolicy.com*). Currently, new ELLs with lower levels of ELP enter the subgroup, while students who attain proficiency in English no longer belong to the subgroup. It is the only subgroup whose composition changes in this way.

Additionally, I recommend that the new iteration of ESEA use the term English language learner or ELL rather than the term limited English proficient students. Just as we do not label first year physics students limited physics students we should not call students in the process of learning English limited-English speakers (LaCelle Peterson & Rivera, 1994).

Now I will discuss **how the common core standards and assessments should address and measure academic outcomes for English language learners**. The new common core standards were developed to provide a "clear and consistent framework to prepare . . . (students) for college and the workforce" (NGA & CCSSO, 2010). While the standards are intended to address all students, ELL experts were not invited to be part of the initial development process. Nonetheless, members of the Working Group on ELL policy and others have since examined the standards and made recommendations regarding how they should be refined to better address the needs of ELLs. These recommendations should be considered and incorporated, as appropriate, into revisions of the common core standards.

With regard to the ELA standards, special attention needs to be given to how and at what point ELLs will be expected to acquire and be assessed in the standards. Because the new common core ELA standards were developed with native English speaking students in mind, it is important to consider the role and use of ELP standards and assessments. Although ELP and ELA are related and even list the same skills (listening, reading, and writing), presumptions about students' background and basic competencies in English differ. Thus, it will be important to articulate the relationship between the two sets of standards and to clearly delineate expectations for when instruction in ELA versus ELP is appropriate for ELLs. This specification should be established in every State or consortium of States by a working committee of ELL and ELA experts using data from current studies of ELA and ELP, as appropriate and available. This committee will need to examine a State's ELP standards and determine at what point along the continuum of learning to speak, read, and write English ELLs at low levels of ELP should be held accountable for ELA standards. This clarification is exceedingly important if States, districts, and schools are to implement and assess the ELA standards in a meaningful way for ELLs as well as for all other students. For ELLs at low levels of ELP, since the ELA continuum starts with the assumption that it is addressing native speakers of English, then it is worth considering substituting the ELP reading and writing standards and assessments as measures of reading and writing achievement for these students.

With regard to mathematics standards, it is important to consider whether these standards need to be addressed only in English or if they can also be addressed in students' native languages. The underlying competencies reflected in the common core standards are benchmarked to international standards and, thus, are based on knowledge and skills that transcend English language proficiency.

Implicit in the national mathematics standards, for example, is the expectation that students can explain methods for solving problems as well as describe, classify, and understand relationships. A crucial factor in meeting these expectations is being able to understand and use the academic language or academic English of different disciplines. While a mastery of academic language is demanding for all students, it can be especially difficult for students who already struggle with other linguistic challenges, such as ELLs and former ELLs.

In a reauthorized ESEA, resources should also be allotted to States to work toward the development of a broad national framework that captures the many dimensions of academic English (Anstrom, DiCerbo, Butler, Katz, Millet, & Rivera, 2010). Currently, the connection between grade-level content goals and the language needed to attain these goals is not made explicit in national or State content stand-

ards. Few educators at either the district or school level have the resources, time or training to perform the kind of linguistic analysis needed to reveal the academic language that creates the most difficulty for ELLs. To this end, The George Washington University Center for Equity & Excellence in Education (GW–CEEE), developed a process to identify the academic language used in assessments, textbooks, and other instructional materials (Anstrom & DiCerbo, in press).

Until a new assessment system is established, it is important for States to continue to work with their existing academic assessments to ensure validity and reliability as well as accessibility to ELLs at different levels of ELP. While many States use accommodations as an approach to make assessments accessible to ELLs, accommodations in the different content areas need to be studied and refined to ensure that they address the linguistic needs of ELLs at basic, intermediate, and advanced levels of levels of ELP. For example, ELLs with basic ELP may benefit more from oral forms of linguistic support and native language support (Pennock-Roman & Rivera, 2010). More research needs to be carried out to determine the most appropriate accommodations, including in ELLs' native languages.

In the interim, States should continue to refine their State assessment policies and communication of those polices to district and school staff responsible for administering State assessments. In the policies, there is great need to clearly distinguish the linguistic needs of ELLs from cognitive, processing, or physical needs of students with disabilities (Shafer Willner, Rivera, & Acosta, 2008). In addition, States need to refine their communication of the policy to district and school staff responsible for administering content assessments so the criteria for administering the assessment and determining appropriate accommodations for individual students are consistent across a State. States should be encouraged to establish and/or improve their systems for monitoring the progress of their ELLs and former ELLs to understand better the relationship of their English language and content knowledge proficiency throughout schooling. Finally, it is important to encourage States to report academic achievement by ELP status and to use these data to make instructional adjustments.

Next I will address **the special challenges faced in developing and implementing assessments which provide information that can support teaching and learning**. The five design principles proposed by NGO/CCSSO in the Common Core Standards hold great promise. It is essential, however, for the learning needs of ELLs, students with disabilities and other special populations to be taken into consideration while the system is being designed, implemented, and evaluated. To address the needs of ELLs, individuals need to be involved who are knowledgeable about second language acquisition, academic English, second language testing, and best practices for instructing second language learners in subject matter content. Equally important, assessments will need to be designed and implemented so ELLs at different levels of English language proficiency are able to access the content of summative, benchmark, and classroom assessments in English.

Development of an integrated learning system implies that, while the goals remain the same, the learning needs of different groups of students must be distinguished and teachers of academic content and teachers of language must be prepared to instruct and assess ELLs at different levels of English language proficiency. A successful system will require retooling of teacher preparation and in-service professional development programs to build the capacity of content and ESL teachers to differentiate instruction and assessment for ELLs, as well as to teach ELLs the academic language they need to access the academic content.

For students in bilingual and dual language situations, it will require teaching and assessing students in the native language as well as in English. For these programs, it is necessary to ensure the content standards and assessments are parallel to the new Common Core Standards.

Every State and consortium should establish an assessment Technical Advisory Committee (TAC) that includes second language testing experts and second language acquisition specialists. The TAC should be responsible for reviewing and commenting on policies, recommending research to be carried out, and providing advice on implementation and refinement of the assessment system.

The delineation of policy at the State and consortium levels is important and should guide practice for the new assessment system. Policies must be developed that clearly define when ELLs are to be included in an assessment, what accommodations are available in English and in the native language for each content area assessed in summative, benchmark, and classroom assessments, and what implementation procedures are to be followed when assessing ELLs at different levels of ELP.

Finally, as part of improving the design of assessments, it is necessary to consider what processes the Department of Education (ED) or other external reviewers will

use to evaluate the new assessment systems. Currently two processes are in place to assess the adequacy of assessments, standards and assessment peer review and title I monitoring, however the processes are not aligned. Whatever review procedures are put in place for the new assessment systems, it is important to ensure the alignment of these processes and that one or more of the individuals involved in a review have knowledge of second language acquisition, language testing, and instruction of ELLs (Shafer Willner, Rivera, & Acosta, 2010).

In conclusion, the design of assessment and accountability systems and their implementation must consider the linguistic diversity and other characteristics of ELLs. To be successful, the system must ensure that the standards and assessment processes addresses academic language as well as English language proficiency. Teacher preparation and in-service professional development programs must build the capacity of content and ESL teachers to differentiate instruction and assessment for ELLs, as well as teach ELLs the academic language required to successfully access the academic content. ELL experts must be involved at every level of design and implementation. States should consider the needs of ELLs in the new standards and assessment system. Policies must address how ELLs are defined, and address procedures for including and accommodating ELLs in summative, benchmark, and classroom assessments. Most importantly the new assessment system must be valid and reliable for all students including ELLs. At the Federal level, ED must improve the review and monitoring of the standards and assessment systems. It is crucial that the review processes explicitly address ELLs and that the reviewers have the necessary expertise and knowledge.

I have great expectations for the ESEA reauthorization and look forward to an interconnected system that works for English language learners.

REFERENCES

Anstrom, K., DiCerbo, P., Butler, F., Katz, A., Millet, J., & Rivera, C. (2010). *A Review of the Literature on Academic English: Implications for K–12 English Language Learners.* Arlington, VA: The George Washington University Center for Equity and Excellence in Education.

Anstrom, K. & DiCerbo, P. (in press). *Final Report: Linking Academic Language to Academic Standards.* Arlington, VA: The George Washington University Center for Equity and Excellence in Education.

LaCelle Peterson, M. & Rivera, C. (1994). Is it real for all kids? A framework for equitable assessment policies for English language learners. *Harvard Educational Review*, 64(1) 55–75.

NGA & CCSSO. (2010) Common Core Standards Initiative. *http://www.core standards.org/*.

Working Group on ELL Policy. (2010). Recommendations for ESEA Reauthorization. *http://www.corestandards.org/*.

Pennock-Roman, M. & Rivera, C. (2010). Mean Effects of Test Accommodations for ELLs and Non-ELLs: A Meta-Analysis. Denver, Colorado: Paper presented at the annual meeting of the American Educational Research Association.

Shafer Willner, L., Rivera, C., and Acosta, B. (2010). *Examination of Peer Review and Title I Monitoring Feedback Regarding the Inclusion and Accommodation of English Language Learners in State Content Assessments.* Arlington, VA: The George Washington University Center for Equity and Excellence in Education.

Shafer Willner, L., Rivera, C., & Acosta, B. (2008). Descriptive analysis of State 2006–2007 content area accommodations policies for English language learners (2008). Prepared for the LEP Partnership, U.S. Department of Education. Arlington, VA: The George Washington University Center for Equity and Excellence in Education. Available: *http://ells.ceee.gwu.edu*.

Short, D. & Fitzsimmons, S. 2007. Double the Work: Challenges and Solutions to Acquiring Language and Academic Literacy for Adolescent English Language Learners. A report commissioned by the Carnegie Corporation of New York. Washington, DC: Alliance for Excellent Education. *http://www.all4ed.org/files/DoubleWork.pdf*.

The CHAIRMAN. Thank you very much, Dr. Rivera.

Now we turn to Dr. Thurlow, National Center on Educational Outcomes from the University of Minnesota.

Welcome. Please proceed.

STATEMENT OF MARTHA THURLOW, Ph.D., DIRECTOR, NATIONAL CENTER ON EDUCATIONAL OUTCOMES, MINNEAPOLIS, MN

Ms. THURLOW. Chairman Harkin, Ranking Member Enzi, and other members of the committee, thank you for asking me to speak today. And I am going to do the best I can because I am suffering from allergies and too much plane riding.

I work with States and districts on the inclusion of students with disabilities in assessments, standards-based reform, accommodations, alternate assessments, and graduation requirements. Today, I want to share what we have learned from research and practice that is relevant to the discussion here.

Students with disabilities who receive special education number 6.6 million students and make up 13 percent of public school enrollment. They are disproportionately poor, minority, and English language learners. The vast majority of them—about 80 to 85 percent—are students without intellectual impairments.

After decades of being excluded from State and district assessments, their participation has now increased to about 97 percent, due in large part to the requirements of ESEA and IDEA. Their academic performance is also increasing. In many cases, students with disabilities have surprised their teachers, their parents, and themselves sometimes by mastering content that before standards-based reform was never taught to them.

Clear, well-defined content standards are the foundation for improved outcomes for all students. The potential benefits of Common Core standards for students with disabilities are great if we can avoid inadvertently developing them in a way that makes it impossible to accurately measure all the content. Clear, well-defined content standards also make it possible to provide appropriate accommodations for students with disabilities, both for instruction and assessments.

We have made tremendous strides in accessible assessments for students with disabilities during the past decade. It is critical that during the development process, we think of all students, clearly define what each assessment is intended to measure and how that content can be measured for all students.

Retrofitting assessments with accommodations and developing a series of alternate assessments because the general assessments do not work for all students is expensive for schools and stigmatizing for students. We also know from title I past practices that out-of-level testing for accountability purposes does not work to improve achievement. It only works to make adults feel better about poor student performance.

As a long-time special educator and assessment expert, I believe that our greatest challenges in improving achievement for students with disabilities are not in the area of the assessments. The greatest challenges are in delivering high-quality instruction in the standards-based curriculum to every student with a disability.

Although there are some ways in which assessments can be improved, unless we provide students with disabilities greater access to the curriculum, making sure that they have individualized instruction, appropriate accommodations, and other supports that they need to succeed, achievement is going to remain low.

With clear and specific standards, teacher capacity to adjust teaching for individual needs can occur without losing the content or performance expectations. These practices will also ensure that other students who are low-performing, predominantly students who are poor and of minority status, but without identified disabilities, also can achieve at higher levels.

The discussion should not be about whether students with disabilities can learn to proficiency as defined for all students, it must be about whether we have the will and commitment to make it happen. We must build on the research that has shown that where there is shared responsibility and collaboration among staff and where students are held to high expectations and are provided specialized instruction, supports, and accommodations so that they meet those high expectations, students can achieve at higher levels and be prepared for college and careers.

It is too easy to explain away the gaps in achievement for students with disabilities by characterizing these students as children to be pitied, who should not be held to the same standards as others because of their disabilities. This characterization is inconsistent with what we know about students with disabilities, and it flies in the face of the purpose of special education. We should expect to see a value-added benefit from the Federal commitment to supplementing State and local funding for special ed services.

This benefit will be realized through the unwavering expectation that all students with disabilities receive high-quality and specialized instruction, have universal access to the challenging grade-level curriculum that is the right of all students, and participate in rigorous and inclusive assessments of their learning.

Thank you.

[The prepared statement of Ms. Thurlow follows:]

PREPARED STATEMENT OF MARTHA L. THURLOW, PH.D.

- Students with disabilities who receive special education number 6.6 million, 13 percent of public school enrollment, and are disproportionately poor, minority, and English-language learners, and the vast majority—about 80–85 percent—are students *without* intellectual impairments.
- The trend lines show increased participation and performance on State assessments—students with disabilities are mastering content that, before standards-based reform, was never taught to them.
- Clear, well-defined content standards are the foundation for improved outcomes for all students and make it possible to provide appropriate accommodations for students with disabilities, both for instruction and assessments.
- We have made tremendous strides in accessible assessments for students with disabilities during the past decade, focusing on the need to think of all students from the beginning of design, clearly define what each assessment is intended to measure, and how that content can be measured for all students.
- Retrofitting is not effective assessment design; practices like out-of-level testing are not effective to improve achievement.
- Our greatest challenges in improving achievement for students with disabilities are NOT in the area of assessments—the greatest challenges are in delivering high quality instruction in the standards-based curriculum to every student with a disability.
- Although there are some ways in which assessments can be improved, unless we provide students with disabilities greater access to the curriculum, making sure that they have individualized, specialized instructions, appropriate accommodations, and other supports they need to succeed, their achievement will remain low.
- With clear and specific, teachable and learnable, measureable, coherent standards, teacher capacity to adjust teaching for individual needs can occur without losing the content or performance expectations.

• These practices will also ensure that other students who are as low-performing as students with disabilities—predominantly students who are poor and of minority status but without identified disabilities—also can achieve at higher levels.

• The discussion should not be about *whether* students with disabilities—or other low-performing students—can learn proficiency as defined for all students—it must be about whether we have the will and commitment to make it happen, building on research that shows it is possible.

• The characterization of students with disabilities as children to be pitied, who should not be held to the same standards as others because of their disabilities, is inconsistent with what we know about students with disabilities—and flies in the face of the purpose of special education.

• We should expect to see a value-added benefit from the Federal commitment to supplementing State and local funding for special education services.

Chairman Harkin, Senator Enzi, and other members of the committee, thank you for inviting me to speak today. I am the Director of the National Center on Educational Outcomes (NCEO), a research and technical assistance organization with funding from the Office of Special Education Programs and the Institute of Education Sciences. NCEO provides assistance to States and districts on the inclusion of students with disabilities in State and district assessments, and on important related topics such as standards-based reform, accommodations, alternate assessments, graduation requirements, universally designed assessments and accessible testing. Because of our focused organizational mission, we work closely with States as they implement standards and assessments for all of their students. We know of the challenges that States and districts face as they work to implement the goals of standards-based reforms. NCEO supports its technical assistance with policy research on current policies and practices in these and other areas. NCEO also conducts other research to move the field forward in its thinking in areas such as how to develop universally-designed assessments that are accessible for students with disabilities without changing the content or level of challenge of the test, and how to most appropriately assess students with disabilities who are also English language learners. We work with other organizations on the critical issues of access to the general curriculum, instruction, and other factors that must be addressed for assessments to show the improved learning that students with disabilities are capable of demonstrating.

I have been a member of the special education professional community since the early 1970s, and have personally viewed the tremendous changes in our country's approach to educating students with disabilities. I have also viewed the stumbles we have made along the way as we determine how to ensure that students with disabilities progress through school and emerge ready for college or a career.

I have been asked to comment on how standards and assessments can be improved to raise outcomes for students with disabilities. I have also been asked to share my thoughts about the special challenges that we face in developing assessments that provide meaningful information about all students. As I address these topics, I want to also make two important points that are critical to understanding the challenges and the promise of standards and assessments for students with disabilities.

IMPROVING STANDARDS AND ASSESSMENTS

To address ways to improve standards and assessments so that they are best for all students, including students with disabilities, it is important to clarify first who students with disabilities are, and also to realize that: (1) students with disabilities have benefited tremendously from our country's focus on standards and assessments, and (2) standards and assessments, by themselves, do not guarantee that student performance will increase, or even that access to the general curriculum and instruction will occur.

Who students with disabilities are. Students with disabilities are not to be pitied or protected from the same high expectations we have for other students. They should not be excluded from the assessments that tell us how we are doing in making sure that they meet those expectations.

Students with disabilities who receive special education as required by the Individuals with Disabilities Education Act currently make up 13 percent of public school enrollment, with percentages in States varying from 10 percent to 19 percent of the State public school enrollment (see Table 1). They are disproportionately poor, minority, and English Language Learners.

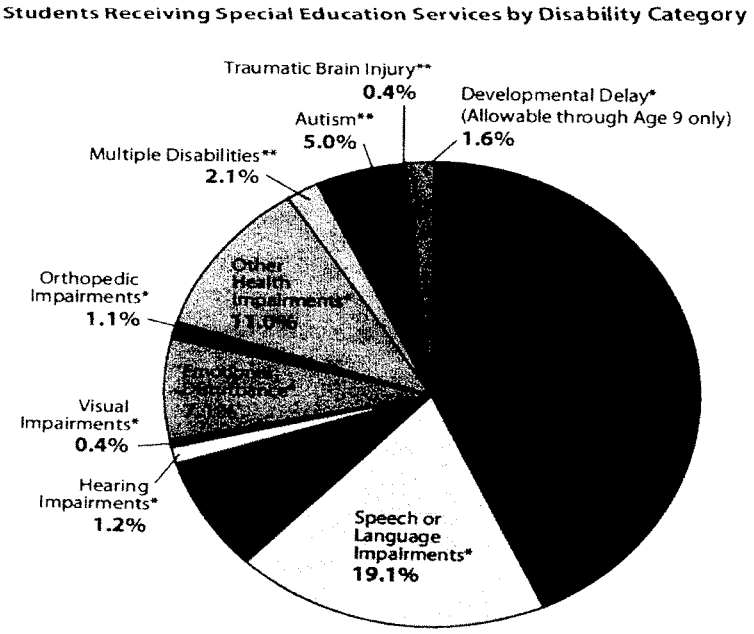
Table 1. Number and Percentage of IDEA Part B Children in Highest and Lowest Percentage States

	No. of Children Served Under IDEA, Part B	Percentage of Public School Enrollment
Highest Percentage States:		
Rhode Island	20,646	19.7
New Jersey	178,870	18.1
Maine	27,987	17.5
Massachusetts	149,743	17.3
Indiana	112,949	17.1
Lowest Percentage States:		
Utah	46,606	10.9
California	468,420	10.6
Colorado	56,336	10.4
Idaho	21,703	10.3
Texas	344,529	10.1
United States Total	6,605,695	13.4

Source: Table 52 of 2009 Digest of Ed Statistics.

The vast majority—about 80–85 percent based on the latest distribution of disability categories—are students without intellectual impairments (see Figure 1). Rather, they are students who with specially designed instruction, appropriate access, supports, and accommodations, as required by IDEA, can meet the same achievement standards as other students. We must ensure that these students progress through school successfully to be ready for college or career. In addition, we have learned that even students with intellectual impairments can do more than we previously believed possible.

FIGURE 1. DISTRIBUTION OF DISABILITY CATEGORIES IN 2008–2009



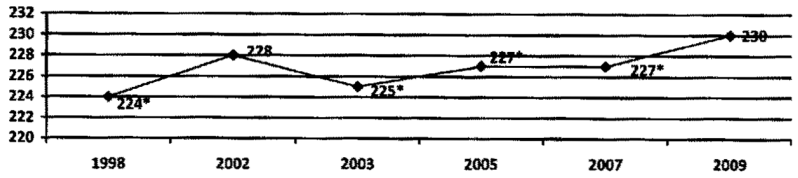
56288-5.eps

In many cases, students have surprised their teachers and parents—and themselves—by mastering content that, before standards-based reform, was never taught to them.

Benefits of standards and assessments for students with disabilities. There is no question that students with disabilities have benefited in many ways from our country's focus on standards and assessments. After decades of being excluded from State and district assessment systems, their participation in State assessments has increased from 10 percent or fewer of most States' students with disabilities participating in the early 1990s, to an average of 99 percent at the elementary level, 98 percent at the middle school level, and 95 percent at the high school level in 2007–2008 (Altman, Thurlow, & Vang, 2010). These increases are due in large part to participation requirements in ESEA and IDEA.

We also are seeing evidence of improvements in the academic performance of students with disabilities. Some of this evidence comes from trends in the performance of students with disabilities on the National Assessment of Educational Progress (see Figure 2 for 2009 grade 8 reading results).

FIGURE 2. NAEP GRADE 8 AVERAGE SCALE SCORES OF STUDENTS WITH DISABILITIES



* = significantly different (p<.05) from 2009

Although there are large gaps in performance between students with disabilities and their peers without disabilities, we have built better understanding about students with disabilities, their opportunities to learn, and what can be expected of them. We have also learned much about what needs to change in their instruction, access to the curriculum, and in assessments in order to first see their achievement increase dramatically, and then to capture that achievement on sensitive assessments.

Standards and assessments do not guarantee improved results or increased access and instruction. Standards and assessment are part of a theory of action that has been driving educational reform in the United States for the past decade or more. It assumes that assessments and accountability promote interventions and improvements in the quality of instruction, which in turn will produce higher performance, which is then rewarded through the accountability system.

This theory of action has been slow to work for several reasons. First and most basic is that current instructional practices, especially for students with disabilities, are not uniformly effective in ensuring success for the students most in need. That is especially true for students with disabilities. Standards and assessments can be improved, but that is no guarantee that the outcomes of students with disabilities will be improved. To raise the outcomes of students with disabilities, we as a nation will need to step up for real change. We must hold our public schools accountable for the learning of students with disabilities, and expect that they commit to practices that we know work. And, given the substantial investment the Federal Government makes annually in support of special education, there need to be better results. We know it is possible because we are seeing success for all students in places with a strong commitment to the learning of all children—all including all students with disabilities. Studies of some of these places have identified what it takes to realize this success:

• In 2004, the Donahue Institute identified 11 practices that existed in such schools, including such factors as: (a) a pervasive emphasis on curriculum alignment with the State standards, (b) effective systems to support curriculum alignment, (c) emphasis on inclusion and access to the curriculum, (d) culture and practices that support high standards and student achievement, (e) well-disciplined academic and social environment, (f) use of assessment data to inform decisionmaking, (g) unified practice supported by targeted professional development, (h) access to resources to

56288-6.eps

support key initiatives, (i) effective staff recruitment, retention, and deployment, (j) flexible leaders and staff that work effectively in a dynamic environment, and (k) effective leadership that is essential to success.

• The National Center for Learning Disabilities (2008) examined successful schools and districts across the Nation, identifying two schools and three school districts where the success of students with disabilities was improved. Though different in location and student demographics, these schools and districts all (a) included students with disabilities in general education classrooms, (b) used data to adjust instruction to each student's needs, (c) changed the ways that general education and special education teachers work together, and (d) restructured administrative organizations and procedures.

• In a recent study of several Ohio school districts where assessment scores showed strong increases over 4 years, Silverman, Hazelwood, and Cronin (2009) found that successful districts shared seven key characteristics: (a) focus on teaching and learning as driver of all decisions, (b) intentional culture shift away from a separate special education model to shared responsibility for all students, eliminating a culture of isolation, (c) collaboration through structures and processes to talk about data and inform instruction, (d) leadership that starts at the district level and uses data to address issues, with monitoring of instructional practice, but shared leadership with principals, building staff, and teacher leaders, (e) instructional practice that ensures access to general curriculum/grade-level content using research-based practices, (f) assessment that includes use of common formative assessments, and (g) curriculum that is aligned, with use of power standards, pacing guides, curriculum calendars, and a relationship to formative assessment.

These three studies, which have looked specifically at what works for students with disabilities, all recognize the importance of standards and assessments. But, they are also about so much more—about the student's access to the curriculum, about a systemwide commitment to all students, and about leadership, collaboration, and shared beliefs among the educators who work with all students, including students with disabilities. Although we can improve standards and assessments, doing so is not a guarantee of raised outcomes for students with disabilities.

Ways to continue to improve standards and assessments. Content standards are the foundation for improved outcomes for all students, including students with disabilities. These standards should identify what students should know and be able to do. Assessments are the means to determine where students are in their knowledge and skills in relation to the standards. A focus on improving standards and assessments should begin by addressing accessibility and universal design. By ***accessibility***, I mean being easy to approach or enter, regardless of barriers that a student might have. Thus, accessible standards are ones that do not have inherent barriers to their attainment, such as a standard that requires a student who is deaf to listen. When I use the term ***universal design***, I refer to a set of principles and procedures that ensure that assessments are appropriate for the widest range of students; universal design techniques can be applied from the beginning of test development to the point when students engage in assessments. The goal of universally designed assessments is to provide more valid inferences about the achievement levels of all students, including students with disabilities.

Improving Standards. Our Nation has recognized the challenges of each State having its own content and achievement standards for students. Those challenges apply to students with disabilities just as they do to students without disabilities. The potential benefits of common core standards for students with disabilities are great. With clear, well-defined content standards, it is possible to better identify appropriate accommodations for students with disabilities, both for instruction and for assessments. And, if we think about all students from the beginning of the development of the common core standards, we can ensure that we do not inadvertently state our standards in a way that makes it impossible to accurately measure their knowledge and skills without instead reflecting their disability. By attending to these concerns from the beginning, we can ensure that rigorous content standards and performance expectations apply to all students, including those with disabilities.

Research evidence on teacher use of accommodations, and accommodations decisionmaking by IEP teams, shows that teachers often have foundational misunderstandings of what the content and achievement standards mean. As a result, strategies to adjust instruction through accommodations often mean that students are denied access to the content; they are either over-accommodated or receive different content than intended by the standards. With clear and specific, teachable and learnable, measureable, coherent standards, teacher capacity to adjust teaching for individual needs can occur without losing the content or performance expectations. Common core standards that are clearer, fewer, and more rigorous should result in

increased clarity for all, assuming that high quality professional development, training, and support continue for ***all teachers*** with ***all students*** as the standards are implemented.

Reading, writing, speaking, and listening standards—given the nature of the standards themselves—often require accommodations for students with disabilities. For example, in the case of students who are deaf, a standard that calls for "listening" should be interpreted to include reading sign language. In a similar vein, "speaking" for some students with speech impairments, for example, should include "communication" or "self-expression." Students who are blind or have low vision should be able to read via braille, screen reader technology, or other assistive technology to demonstrate their comprehension skills. "Writing" should not preclude the use of a scribe, computer, or speech-to-text technology for students with disabilities that interfere with putting pen to paper, for example.

Assessments. We have made tremendous strides in making assessments more accessible for students with disabilities during the past decade. States and test developers have, in general, started the development of their assessments with the recognition that students with disabilities are general education students first. The implication of this is that assessments are better designed from the beginning with all students in mind, and should not preclude the participation of most students with disabilities. It is critical that during the development process we think of all students, clearly define what each assessment is intended to measure, and how that content can be measured for all students. Retrofitting assessments with accommodations and developing a series of alternate assessments because the general assessments do not work for all students is expensive for schools and stigmatizing for students.

The research base for developing accountability assessments that are more appropriate for all students has dramatically increased in the past several years. Based on this research, NCEO developed five principles for assessments used for accountability (Thurlow ET al., 2008):

- All students are included in assessments in ways that hold schools accountable for their learning.
- Assessments allow all students to show their knowledge and skills on the same challenging content.
- High quality decisionmaking determines how students participate.
- Public reporting includes the assessment results of all students.
- Accountability determinations are affected in the same way by all students.
- Continuous improvement, monitoring, and training ensure the quality of the overall system.

Each of these is supported by specific characteristics of assessment systems that are appropriate for all students, including students with disabilities. All together, they provide an important framework for any future assessment system.

These principles reinforce what we have learned—first, thinking about students when assessments are first designed, developed, and implemented; second, defining allowable accommodations as part of the development process; and third, ensuring that the assessment system include all students, without exception. This way, developers have focused on ensuring that tests really measure what they are intended to measure—not extraneous factors, such as whether the students can figure out what the test developer means by a question or whether a picture has important clues about the answer to a question (Dolan ET al., 2009; Thurlow ET al., 2008; Thurlow ET al., 2009). Identifying ways to improve assessments for students with disabilities has, in fact, resulted in improving assessments for all students.

What these principles do not do is indicate the specific nature of the assessment. Whatever the assessment approach—computer-based assessments, through course assessments, or paper and pencil end of course assessments—the critical point is to think about the whole population of students, including students with disabilities. Taking computer-based assessments as an example—these assessments show promise for increasing the accessibility of assessments. They also make it easier to fall back into some pitfalls that have been demonstrated to create problems for the assessment of students with disabilities. On the positive side, computer-based assessments can be developed in a way that embeds what are called "accommodations" when the test is paper-based, such as the following described by Russell (2008):

- Users navigate and interact with the functional elements of the test delivery system using a standard mouse, keyboard, touch screen, intellikeys, switch mechanism, sip-and-puff device, eagle-eyes, and other assistive communication devices.
- Text can be read aloud using a human voice or a synthesized voice, or can be signed.

• All graphics, drawings, tables, functions, formulas, and other non-text-based elements of an item can be provided through spoken descriptions.

An auditory calming tool can be provided that allows all students to select from among a list of pre-approved sound files, and play softly in the background as the user works on the test. A computer-based system could record each use of an incorporated feature or accommodation to document use for individual items as well as overall. There are tremendous possibilities for dramatically increasing the accessibility of assessments in a computer-based assessment system based on grade-level content standards. These assessments also have the potential to aid teachers as they determine how to move students to grade-level achievement.

Computer-based systems also make it easier to fall back into some pitfalls that have been demonstrated to create problems for the assessment of students with disabilities. We must avoid pitfalls of the past in designing computer-based systems. They should be developed to be as transparent as possible about the content on which students are assessed and the ways in which the content is assessed. They should not revert to normative assessments, which compare students only to each other rather than to content standards, even in the name of being able to measure growth. Title I evaluation systems prior to 1994 were based on these types of approaches, and demonstrated dramatically that schools can show that students make "progress," but the progress is meaningless if it is not tied to the intended content and achievement targets. These practices resulted in the failure of the system in identifying where schools were succeeding and where they were not. Students remained far behind their peers—and even increased the achievement gaps—in schools deemed successful based on flawed testing assumptions. Computer-based systems should not revert to an out-of-level testing approach. To avoid the mistakes of the past, any adaptive computer-based assessments must be on grade-level. Even when constrained to grade-level, adaptive testing practices must be transparent enough to detect when a student is inaccurately measured because of splinter skills common for some students with disabilities, for example, with poor basic skills in areas like computation and decoding, but with good higher level skills, such as problem solving, built with appropriate accommodations to address the barriers of poor basic skills.

The research base has dramatically increased for new forms of assessments, like alternate assessment based on alternate achievement standards (AA–AAS), developed to measure the academic achievement of a very small number of students who have the most significant cognitive disabilities. NCEO, in collaboration with the National Alternate Assessment Center (NAAC) has conducted an extensive literature review and has identified 10 common misperceptions about AA–AAS, as well as research-based recommendations to ensure common understanding and high quality assessments (Quenemoen, Kearns, Quenemoen, Flowers, & Kleinert, 2010). A summary of the research-based recommendations is included in Appendix A.

CHALLENGES IN PROMOTING IMPROVED ACHIEVEMENT FOR STUDENTS WITH DISABILITIES

Our greatest challenges in improving achievement for students with disabilities are NOT in the area of assessments. Including all students in assessment and accountability systems as well as requiring reporting of assessment results broken out by student groups that historically underperform has been critical in helping us understand our great challenges. These greatest challenges are in delivering high quality instruction in the standards-based curriculum to every student with a disability. Although there are some ways in which assessments can be improved, the real work that needs to be done is in providing students with disabilities greater access to the curriculum, making sure that they have the individualized instruction required by IDEA as well as appropriate accommodations and other supports they need to succeed. States that have done this have seen the improved results.

We know how to educate all children, including those with disabilities, if we have the will to do so. The discussion should not be about *whether* students with disabilities can learn to proficiency—and thus, it should not be about *whether* they should be included in the assessment and accountability measures we have for all students—it must be about whether we have the will and commitment to make it happen. We must build on the research that has shown that where there is shared responsibility and collaboration among staff, and where students are held to high expectations and are provided specialized instruction, supports and accommodations so that they can meet those high expectations, students score higher on assessments.

Still, there are some risks as we move forward to develop assessments based on common core standards. It is too easy to explain away the gaps in achievement for

students with disabilities by characterizing these students as poor little children who should not be held to the same standards as others because of their disabling condition. This characterization is inconsistent with what we know about students with disabilities—and flies in the face of the purpose of special education. We should expect to see a value-added benefit from the Federal commitment to supplementing State and local funding for special education services. This benefit will be realized through the unwavering expectation that all students with disabilities receive high quality and specialized instruction, have universal access to the challenging grade-level curriculum that is the right of all students, and participate in rigorous and inclusive assessments of their learning.

Thank you.

REFERENCES

Altman, J., Thurlow, M., & Vang, M. (2010). *Annual performance report: 2007–2008 State assessment data.* Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.

Dolan, R.P., Burling, K.S., Harms, M., Beck, R., Hanna, E., Jude, J., Murray, E.A., Rose, D.H., & Way, W. (2009). *Universal design for computer-based testing guidelines.* Iowa City, IA: Pearson.

Donahue Institute (2004), *A study of MCAS achievement and promising practices in urban special education.* Hadley, MA: University of Massachusetts Donahue Institute.

National Center for Learning Disabilities. (2008). *Challenging change: How schools and districts are improving the performance of special education students.* New York: Author.

Quenemoen, R., Kearns, J., Quenemoen, M., Flowers, C., & Kleinert, H. (2010). *Common misperceptions and research-based recommendations for alternate assessment based on alternate achievement standards* (Synthesis Report 73). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.

Russell, M. (2008). *Universal design of computer-based tests: RFP language.* Unpublished document, Boston College.

Silverman, S.K., Hazelwood, C., & Cronin, P. (2009). *Universal education: Principles and practices for advancing achievement of students with disabilities.* Columbus, OH: Ohio Department of Education.

Thurlow, M.L., Quenemoen, R.F., Lazarus, S.S., Moen, R.E., Johnstone, C.J., Liu, K.K., Christensen, L.L., Albus, D.A., & Altman, J. (2008). *A principled approach to accountability assessments for students with disabilities* (Synthesis Report 70). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.

Thurlow, M.L., Laitusis, C.C., Dillon, D.R., Cook, L.L., Moen, R.E., Abedi, J., & O'Brien, D.G. (2009). *Accessibility principles for reading assessments.* Minneapolis, MN: National Accessible Reading Assessment Projects.

APPENDIX A: RETHINKING ASSUMPTIONS ABOUT ALTERNATE ASSESSMENT BASED ON ALTERNATE ACHIEVEMENT STANDARDS

To facilitate the process of rethinking assumptions about alternate assessments based on alternate achievement standards (AA–AAS), common misperceptions are identified first, followed by the assumptions underlying them and a research response to those assumptions. A comprehensive summary of the literature underlying the research responses is provided in ***Common Misperceptions and Research-based Recommendations for Alternate Assessment based on Alternate Achievement Standards*** *(NCEO Synthesis Report 73 by Quenemoen, Kearns, Quenemoen, Flowers, & Kleinert).*

Common misperception No. 1.—Many students who take the AA–AAS function more like infants or toddlers than their actual age, so it makes no sense for schools to be held accountable for their academic performance.

- **Assumptions Underlying Misperception:** Some people assume that students who take the AA–AAS have such severe disabilities that they are unable to learn academic content. Sometimes, this misperception is rooted in the assumption that all students must progress through typical infant and preschool skill development before any other academic instruction can occur.
- **Research Response:** First, learner characteristics data from many States show us that MOST students who participate in AA–AAS have basic literacy and numeracy skills. Second, we have understood for many decades that waiting until these students are "ready" by mastering all earlier skills means they "never" will be given access to the skills and knowledge we now know they can learn. In the

1980s, educators realized that students with significant disabilities could learn functional skills to prepare for independent adult life, even before mastering all lower skills. In recent years, research suggests that these students can often also learn age-appropriate academic skills and knowledge even when they have not mastered all earlier academic content.

Research-based Recommendation: Build accountability systems to ensure that all students who are eligible for the AA–AAS have access to and learn academic content expected for their same-age typical peers, to an appropriate but challenging alternate achievement standard.

Common Misperception No. 2.—Many students who participate in AA–AAS have life-threatening medical conditions or are not able to communicate.

• **Assumptions Underlying Misperception:** People sometimes assume that AA–AAS students are a small homogeneous group of students with multiple problems that go well beyond what schools can actually handle; these students cannot speak, hear, or communicate in any way.

• **Research Response:** Students who participate in AA–AAS are generally less than 1 percent of the total student population or about 9 percent of all students with disabilities. Most of the students who take the AA–AAS (90 percent) have consistent communication skills. Only about 10 percent of AA–AAS students communicate on a pre-symbolic level (without intentional use of language, pictures, objects, or signs). These students *can* communicate, but need to be given opportunities to learn effective strategies, including the use of assistive devices.

Research-based Recommendation: For the small group of students who initially demonstrate a lack of symbolic communication (about 10 percent of students who take the AA–AAS), educators should persistently and systematically seek multiple and varied communication strategies including assistive technology to permit these students to learn and then to show what they know on an AA–AAS.

Common Misperception No. 3: Students in the AA–AAS can learn only rote academic skills, so AA–AAS should reflect only these skills.

• **Assumptions Underlying Misperception:** People sometimes assume that the curriculum for students with severe disabilities often has been based on math skills of time and money and reading skills limited to sight words because that is all these students can learn.

• **Research Response:** It is true that research through the 1990s reflects a very narrow curriculum. Researchers now are finding strong evidence of academic skills and knowledge development among these students, including abstract concepts and transfer of learning, for students who participate in AA–AAS. We are only beginning to learn what these students are capable of, once given the opportunity to learn and access to appropriate accommodations such as assistive technology. In our work with States, we have encountered many teachers who have been "surprised and amazed" at what their students are able to learn when given the chance.

Research-based Recommendation: Build AA–AAS approaches based on a model of academic content development that allows these students to demonstrate a range of grade-level content that their peers are also learning and demonstrating.

Common Misperception No. 4—The AA–AAS has eliminated the teaching of important functional skills.

• **Assumptions Underlying Misperception:** People sometimes assume that the addition of academics to the curriculum for students with severe disabilities means that there is limited time for teaching functional skills like self-care, community participation, and safety. There is not enough time in the day to do both.

• **Research Response:** AA–AAS are designed to ensure students with significant cognitive disabilities are taught academic content like their peers, but a student's IEP will often still include important functional skill goals. Many teachers have found that blended instruction in academic and functional skills yields better results for both. The "line" between academics and functional instruction begins to blur as teachers and parents discover how truly useful and satisfying increased literacy and numeracy skills are for these students, for quality of life and enjoyment, for integration into the community, school, or adult life, and for future employment.

Research-based Recommendation: Provide training and support to teachers so that they can effectively merge academic and functional instructions where appropriate and so that they understand the vital importance of academic skills and knowledge to full participation in family, school, and community life.

Common Misperception No. 5—AA–AAS must cover all of the same content that is on the general assessment for typical peers.

• **Assumptions Underlying Misperception:** People sometimes assume that the grade-level curriculum is very challenging and has far too much information for these students to cover in a year, let alone learn at all, but Federal law requires the same content on all tests.

• **Research Response:** Federal regulations permit States to define the appropriate depth, breadth, and complexity of content coverage for the AA–AAS. Researchers are working on ways that students can access grade-level content at various "entry points." States must show that these content priorities truly "raise the bar" of historically low expectations, and are clearly linked to the content that typical students in the same grade should know and be able to do. Since this is a shift for teachers who do not have experience with this content, training and support to teachers is an essential component of high quality alternate assessments.

Research-based Recommendation: Provide training to teachers, and to other key assessment system stakeholders and advisors, on what research suggests these students are able to know and do when given the opportunity.

Common Misperception No. 6—Most AA–AAS are entirely individualized and differ for each student.

• **Assumptions Underlying Misperception:** People sometimes assume that teachers make so much adaptation and adjustment to the assessment for each student that there is no way you can compare results from one school to another.

• **Research Response:** A good AA–AAS allows a defined amount of flexibility in administration of the items and tasks because students with the most significant cognitive disabilities vary in how they take in and respond to information and requests. Even so, AA–AAS must also adhere to basic standards of technical quality so that the scores can be compared for accountability purposes. An AA–AAS should incorporate training, oversight, and structures to balance flexibility with standardization of procedures and ongoing monitoring to ensure the assessments are administered, scored, and reported as intended.

Research-based Recommendation: All AA–AAS scores should indicate whether the student is proficient in an academic domain through procedures that allow flexibility but control for possible sources of error.

Common Misperception No. 7—An AA–AAS measures teacher performance in compiling attractive portfolios or examples rather than measuring student academic performance.

• **Assumptions Underlying Misperception:** People sometimes assume that teachers who are able to put together pretty portfolios or examples, or who can choose student examples that make them look good, will score higher than teachers who may teach well but who do not spend time creating pretty portfolios or examples of what their students do.

• **Research Response:** Given what we understand about student characteristics, most AA–AAS formats require test administrators familiar to the student. That means that in most cases, teachers interact with the student to capture accurate evidence of what the student knows and can do. A good AA–AAS is designed to control for administrative responses that are decorative, and to focus on independent student performance. Research has shown that teachers who are well-trained in instruction and assessment administration often have students with higher AA–AAS scores, but spending a lot of time making the portfolio "look good" has little impact on scores.

Research-based Recommendation: Train teachers on systematic data gathering procedures, provide oversight, coaching, and monitoring to ensure they implement the procedures as intended, and design scoring processes to exclude evidence that reflects teacher behaviors instead of independent student performance.

Common Misperception No. 8—It would make more sense if teachers simply reported on their students' progress meeting IEP goals rather than requiring an AA–AAS.

• **Assumptions Underlying Misperception:** People sometimes assume that students with the most significant cognitive disabilities have IEPs that define what they should be learning. Gathering data that already are used for the IEP is the best measure of the students' achievement.

• **Research Response:** A good IEP will identify the services, supports, and specialized instruction needed so that the student can learn both academic and functional skills and knowledge. Data gathered on the specific goals and objectives in the IEP are important for *individual accountability* among IEP team members for these short- and long-term goals and objectives, in all areas where the student has them. Some of these goals and objectives will specify the services and supports the

student needs to *access* the general curriculum, but student progress based on the IEP does not provide accountability for student *achievement* of proficiency in the general curriculum. In contrast, AA–AAS are designed to provide data for *system accountability* to ensure that all students are provided access to and are achieving to proficiency in the general curriculum.

Research-based Recommendation: Design AA–AAS so that there are comparable data on the effectiveness of schools in providing access to the general curriculum to students with the most significant cognitive disabilities.

Common Misperception No. 9—Some AA–AAS formats (i.e., portfolio, checklist, performance assessment) are better than others.

• **Assumptions Underlying Misperception:** People sometimes assume that one method is better than another, with "better" meaning more technically adequate; the specific method that is considered better or worse often is based on good or bad experiences in the past.

• **Research Response:** Research on the technical quality of AA–AAS has shown that the format of the test is a poor predictor of technical quality. What a "portfolio" or "checklist" or "performance assessment" or what any other type of format name is can vary enormously, and a number of States now use hybrid models that combine elements of these approaches. Any of these types of formats can be of poor or high quality. A good AA–AAS should sensitively and accurately measure what students know and can do once they have been given appropriate access to interesting, age-appropriate academic content.

Research-based Recommendation: Select the format of the AA–AAS based on beliefs about academic teaching and learning for AA–AAS students.

Common Misperception No. 10—No AA–AAS can be a technically adequate measure of student achievement for accountability purposes.

• **Assumptions Underlying Misperception:** People sometimes assume that the AA–AAS breaks all the rules of good design of large-scale assessments as judged by high quality psychometric evidence that have been used by measurement experts for a century.

• **Research Response:** The challenges of designing AA–AAS are very new; prior to the 1990s, no large-scale assessment program included students with significant cognitive disabilities, and very few measurement experts had experience designing assessment for these students. Fortunately, there has been a great deal of work done since the 1990s on issues that have emerged in developing psychometrically sound AA–AAS. AA–AAS can be designed to produce valid and reliable information about student outcomes.

Research-based Recommendation: State assessment offices should address three components of the assessment design as they develop and implement the AA–AAS: (a) description of the student population and a theory of learning for these students, (b) structure of the observations from the assessment, and (c) interpretation of the results. The technical defense of an AA–AAS starts and ends with these three components.

The CHAIRMAN. Thank you very much, Dr. Thurlow.

Thank you all for your testimonies, and we will start a round of 5-minute questions here.

Dr. Thurlow, you know probably of my involvement with the whole disability movement for many, many years. This is one area that I intend to focus on a lot in the reauthorization of ESEA. What steps can we take to ensure that the assessments we develop are appropriate for all students, including students with disabilities, and provide us with a valid and valuable information on their achievement and growth?

What can we learn from I think it was what Dr. Phillips talked about? In other words, a technology-based, computer-based system that gives perhaps a more rapid and more thorough information to teachers on how to assess their students, and especially students with disabilities.

A subset of my question might be you are familiar with the 1 percent, 2 percent problem? The exemptions for the 1 percent. Now

they want to go to 2 percent. If you could address yourself to that briefly, I would appreciate that.

Ms. THURLOW. All right. Let me start with steps to take, and I think we have been learning this across time as we have worked with State assessments, that we need to take that universal design approach, where we start from the very beginning, thinking about all students. That means when we are talking about our standards that we be clear about exactly what they mean so that we know what accommodations can be provided that won't get in the way of what we are trying to measure.

We need to think about those accommodations carefully so that we are getting valid measures. We need to continue to work in relation to that on the decisionmaking process so that students are not over accommodated, for example, which in many cases ends up interfering with their performance.

Well, let me jump to the notion of technology-based. In fact, I would broaden that to the variety of discussions we are having about innovative assessments. Most of them are going to be wonderful for children with disabilities. It is not the particular approach we take. It is how we ensure that we have thought about all students as we take a particular approach.

Talking about technology-based assessments, I think it has tremendous potential in being able to incorporate what we now call accommodations. They don't have to be separate. It can be part of the assessment itself. That is a big advantage, and all the others, getting scores quickly, etc, are advantages.

One caution I would have is that we need to remain on grade level. We need to continue the same expectations for students with disabilities as we have for other students. Dr. Phillips talked about every student getting the same content standards. We need to make sure that that happens for students with disabilities as well, that we don't somehow send them down a path where they don't get to all of the standards that everybody else gets to.

The CHAIRMAN. One percent. We had the 1 percent exemption, and now people are pushing for 2 percent.

Ms. THURLOW. OK, I would never call——

The CHAIRMAN. Of course, we know that 1 percent translates into 10 percent.

Ms. THURLOW. Ten percent of students with disabilities, 1 percent of the total population. That is like a general estimate.

The CHAIRMAN. That is right.

Ms. THURLOW. The alternate assessment based on alternate achievement standards, which we typically refer to as the "1 percent assessment," or often refer to that, I think has been a tremendous benefit for the field. We have figured out who the students are, pretty much, who belong in that assessment, those students with significant cognitive disabilities, intellectual involvement. We have made tremendous strides in figuring out what the content standards are, how they apply to those students, and we are working and evolving in our knowledge of how best to assess those students.

Remember, these students were never in assessments before. We have made tremendous improvement there. I believe the 1 percent is pretty good, pretty accurate percentage for students to be in-

volved in the alternate, based on the alternative achievement standards.

The CHAIRMAN. What about 2 percent?

Ms. THURLOW. I think 2 percent, this is the alternative assessment-based, a modified achievement standard, a relatively new assessment. We are really looking at who those kids are. There have been challenges in identifying what makes students with disabilities different from other low-performing students. So we see the same characteristics.

They are generally poor students, low-performing——

The CHAIRMAN. While I found that there was maybe some acceptance among the disability community for the 1 percent, I find almost no acceptance for the 2 percent. And it just goes too far.

Ms. THURLOW. It has become controversial, holding different standards for another 2 percent.

The CHAIRMAN. I think we have to look at that very carefully.

Ms. THURLOW. Thank you.

The CHAIRMAN. Well, my time has run out. I have other questions for Dr. Phillips, too, on assessments. But, I will do that in the next round.

Senator Enzi.

Senator ENZI. Thank you, Mr. Chairman.

I appreciate the testimony of all these witnesses, and I have a lot of questions, too.

I will begin with Dr. Paine. How can the Federal Government support the work of the States in the further development and adoption of the Common Core standards without nationalizing or federalizing the standards?

Mr. PAINE. Great question. I certainly appreciate the question.

I think the revised ESEA should reward State leadership and innovation not just with funding for assessments, and professional development and other inputs, but also by perhaps codifying a new State-Federal partnership that does, in fact, promote innovation in alignment of practice to this set of Common Core standards. By that, I mean allowing States some degree of flexibility in establishing an accountability system that works for that particular State.

I think also that as we look at the innovation money that is available right now, the $350 million Race to the Top innovation money that is available, there have been two assessment consortia, if you will, that CCSSO is going to work with. One is a fairly traditional-based summative assessment approach with some degree of balance, and the other really includes multiple measures of looking at how we assess student progress.

And so, I think that a Federal role could be recognizing that if we really truly are interested in 21st century types of assessments that really will link kids to the workforce, to the private sector—our own John Chambers from Cisco hails from Charleston, WV. And so, as we engage in conversations with John, he clearly says that it is about kids knowing content at a high level and a proficient level, but it is also about kids understanding how to apply that content.

You simply don't measure that kind of performance result necessarily with a summative standardized type of test. Looking at

adaptive tests and innovative tests, at ways to assess student progress in many innovative, different ways so that you are measuring the full scope of these rich, robust common State standards.

That is certainly a role that the Federal Government could play and Congress could play in the reauthorization in encouraging those kinds of innovative assessment systems with strong accountability measures.

Senator ENZI. Thank you.

Dr. Schmeiser, given the work that ACT has done with WorkKeys, could you explain how that would relate to the issue of career-ready standards? How would that be assessed?

Ms. SCHMEISER. Yes, Senator.

Matter of fact, the WorkKeys program, which is a workforce development program offered by ACT, has been predicated on a database of over 17,000 jobs in the United States that have been profiled. That data fed right into the Common Core development process. As I mentioned, that was very much an evidentiary, research-based process.

We used information about what is needed for high school graduates. What do they need to know and be able to do when they leave high school in order to be able to go into workforce training programs and be ready to learn job-specific skills?

That information and data on those foundational skills fed into that evidentiary base in being able to define the Common Core. When we talk about college and career ready, the career-ready evidence from WorkKeys was used as part of that process. It has been front and center in the evidence that was used to identify the Common Core, and WorkKeys will, in fact, be aligned with the Common Core as well.

Senator ENZI. Thank you.

Dr. Phillips, it is clear that States need to update their State-wide assessments to align with these new State-developed standards. Will State-wide summative assessments provide accurate assessment of the student knowledge of these standards, or will additional assessments be needed?

Mr. PHILLIPS. Well, I think the plan with the consortia of States and the common standards, assuming that those are adopted, that would go a long way toward solving the problem—instead of going in 50 different directions, they might go in 2.

Assuming that they also are able to set high internationally benchmarked performance standards on, let us say, both consortia or however many there might be, then that should go a long way toward solving the problem that I discussed.

Senator ENZI. Thank you.

I can see that my time is about up. So I will save some for the next round.

The CHAIRMAN. Thank you, Senator Enzi.

As Senators know, we sent around to your offices that, with the concurrence of Senator Enzi, we have adopted a new procedure here in this committee that the Chair will recognize Senators in the order of their appearance at the committee dais. I think that is just a more fair and just way of doing things.

The order I have would be Senator Murray, Senator Alexander, Senator Franken, Senator Isakson, and then Senator Bennet, Senator Hagan, Senator Merkley, and Senator Casey thus far.

With that, I would then now recognize Senator Murray.

SENATOR MURRAY

Senator MURRAY. Thank you very much, Mr. Chairman.

Thank you to you and Senator Enzi for holding this really important hearing in this series, and I really appreciate all of the witnesses today.

I am very interested in making sure all of our students succeed. I think we all are. Dr. Schmeiser, you talked a little bit about making sure that a student is prepared for a college or a career. And I am interested, as we see the dropout rate so high today and a lot of our students not succeeding, if you can talk a little bit—any of you—about how we can prepare students for both post-secondary education and a career.

What are differences, if there are any, in the skills that a student needs to be successful in a post-secondary education program or in a workplace, and what is it like, and how do we write an assessment that makes sure that all kids fall into a category of success no matter where they intend to go? I will open it up to anybody who would like to respond.

Ms. SCHMEISER. Thank you, Senator Murray.

I would like to say that when the Common Core State Standards Initiative got underway, the definition that they arrived at for college and career readiness assumed that students, all students could be educated to a common standard, so that when they leave high school, they are ready to go into some form of post-secondary, whether it is 2-year, 4-year, trade, technical school, or go into workforce training programs for the kinds of jobs that I described in my testimony.

The purpose of the standards is to set a common expectation for all students when they leave high school so they are ready to go ahead and go into post-secondary without needing remediation or go on to workforce training programs and learn the job-specific skills that they will need in their career, as well as some of the nonacademic behavior, the good job behaviors that go along with that.

Senator MURRAY. I hear what you are saying is that we can do it, but I am asking you what is that? What is it that we are doing that says that we have an assessment that reaches both?

Ms. SCHMEISER. That the assessment reaches both?

Senator MURRAY. What do we need to do in our high schools different today than we have been doing that makes sure our students reach both of those potentials?

Ms. SCHMEISER. I think what we need to do, it goes back to needing an aligned system that not only talks about Common Core standards, but also has aligned professional development for teachers so they understand what the standards are and what they mean. They can teach those standards in many different ways.

The idea is not that one-size-fits-all in the instructional process, but those standards can be contextualized in career formats. They can be also introduced in more academic formats. The point is, the

system has to be aligned both in terms of outcome, instruction, assessment, and the data systems coming back so they can identify when students are falling behind, whether it is in a career contextualized course or an academic course.

Bottom line, all kids are educated to the same standards.

Senator MURRAY. Anybody else want to comment on that?

Dr. Rivera.

Ms. RIVERA. I would just like to say that if we are considering the role or what should happen for English language learners and other students, different populations, it is very important that teachers know how to translate those standards and that they are able to address the individual differences of those subgroups of students.

For English language learners, I really believe, and I believe for other students as well, not just English language learners, that this whole idea of academic language is critical. And that teachers need to understand what—to dissect the standards and actually understand the language of the content and be able to teach it explicitly to the students, and this will work for English language learners. It will work for many different subgroups of students.

In other words, that language includes—and I know it may sound like we have—it is the English language. But, yes, we have to teach students how to understand the phrases, the language of academics, the language of the classroom, as well as the specific language of the content. In biology or wherever.

I will just give you an example. We have been working with California and with New York. In New York, we are using the biology, we are working with them to help identify the language of biology and then to teach teachers how to explicitly teach that language so that those students can then be successful in that content area. This is the kind of work that needs to be done for all students really.

Senator MURRAY. OK. Thank you.

Thank you, Mr. Chairman. My time is up, but I do want to be able to submit questions. I assume we can do that for the record.

The CHAIRMAN. We will leave the record open for questions. No doubt about that.

Senator Alexander.

STATEMENT OF SENATOR ALEXANDER

Senator ALEXANDER. Thanks, Mr. Chairman. And thanks for an excellent hearing and excellent witnesses.

If you will permit me a little historical context? I went back and found an article from 1991, entitled "What We Were Doing When We Were Interrupted." The "we" was we in the George Bush I Education Department. "Interrupted" meant the election in 1992. What it reminded me was that of the national goals that Senator Harkin mentioned that in 1987, the Governor and the President setting national goals, and then President Bush's America 2000 strategy to mobilize the country to meet those goals.

These initiatives included, No. 1, a new set of national standards in core curriculum subjects, including science, history, English, geography, art, civics, and foreign languages. Math was already done.

No. 2, a voluntary national examination system geared to those new standards.

Then when we left, there were, according to this, seven task forces created to develop new national academic standards are funded and scheduled to complete their work by 1994, 1995. Some of us were discussing earlier there was a Goals 2000 panel that was going to push that forward.

Now, I compliment the work that the States have done so far, and I am watching it cautiously. English and math are the easy parts of a very hard thing to do. I remember the history standards in the 1990s. They completely blew up, and I want to see how you do this with U.S. history when the time comes.

I guess my first question, and I will ask you, Dr. Schmeiser, just give, if you can, a fairly short answer. College and career ready, do you mean college or career ready? Do you mean to say every student should go to college?

Ms. SCHMEISER. No. I think the point was whether a student goes on to college after high school or into a workforce training program, they will be educated to the same standards, not different standards.

Senator ALEXANDER. Well, I mean, how realistic is—how many today go to college of our high school graduates? What percent do—half, 60 percent, 40 percent?

Ms. SCHMEISER. Well, I think there are estimates that up to three-quarters of our Nation's high school graduates go into some form of post-secondary education within 2 years of leaving high school.

Senator ALEXANDER. Yes, and many aren't prepared.

Ms. SCHMEISER. Too many are not prepared. Yes, sir.

Senator ALEXANDER. I am interested in what you have found. Dr. Phillips, you mentioned going in 2 directions instead of 50. In conversations I have had with Secretary Duncan, I worried a little bit about—you know, I have been interested to see how the common standards worked.

The tension that happened in the 1990s, as I remember it, was—and going back to Senator Enzi's point, I think there is a difference between national and Federal. National to me means States getting together, doing things. That is national in our very diverse constitutional system, which is very different than Taiwan and Singapore—small, people very much the same. Federal means Washington meddling in that.

I had wondered whether it might not be even a good idea if, as things went along, we might have two or three common, maybe a Massachusetts-led coalition of States, maybe an Iowa-led coalition of States. I believe you were talking about maybe one type of assessment and another type of assessment.

In other words, to build into this effort to raise standards enough diversity to provide a safety valve, which is a safety valve against mediocrity, for one thing, to make sure that national doesn't mean average. To avoid political correctness or the feeling of one part of the country having a view imposed on it that it doesn't agree with, say, as history standards or other standards come up.

What has been the thinking on this as you all have worked through this?

Mr. PHILLIPS. Well, I think that is right. I don't know how many consortia of States will be funded, but that is about the right number. And what is important is this is a substantial improvement over what we have today, where each State is going in a different direction.

Senator ALEXANDER. Right.

Mr. PHILLIPS. One thing I would like to say is that these are grassroots efforts, and Federal funding of these efforts doesn't make it Federal. I believe these efforts ultimately will need Federal funding. There are many examples where the Federal Government has provided funds without being in charge or in control.

Senator ALEXANDER. I can agree with that, although one has to be careful, as we move on. Mr. Chairman, on his point about Tennessee standards, you are exactly right. They were low. I always thought the cure for that, though, was just to establish a rating system and say, like you do in football, there is Division 1, 2, and 3, and just tell the people of Tennessee, they are playing in Division 3, and they would very quickly be embarrassed into Division 1.

They have actually changed under the Governor's leadership and partly just because of the embarrassment of what you just described.

Thank you, Mr. Chairman.

The CHAIRMAN. Thank you, Senator Alexander.

Senator Franken.

SENATOR FRANKEN

Senator FRANKEN. Thank you, Mr. Chairman.

Senator Alexander, I think, was talking in regards to some of the questions I have, which is sort of how to set national standards and make sure they are national. I understand the importance of State flexibility and local flexibility beyond State flexibility in implementing these standards, but Dr. Phillips, I was kind of concerned. I mean, you did give the examples of the loophole in NCLB, which is to set these very low standards in some States.

I am wondering how we are able to have Common Core standards, but are States, how do we guard against States still using that loophole?

Mr. PHILLIPS. Well, one thing would be to be aware of this when ESEA is reauthorized. Therefore, it is on your radar.

Senator FRANKEN. That is what we are doing.

Mr. PHILLIPS. Right. Exactly.

Senator FRANKEN. Right. Right.

[Laughter.]

Mr. PHILLIPS. One way you could do this is if you adopt common standards, you have closed part of the loophole. What you have done there is you now have a reasonably common set of skills that you want students to learn. What you then need to do is take the second step and rein in these discrepancies in the performance standards. It is the cut scores on the tests.

Many States that have low-performance standards have challenging content standards. In other words, they tell the press we are expecting all this of our students, but then they lower their cut score so that all the students pass. One way around that is through

international benchmarking, where I am assuming if you have three or four consortia, when they set that cut score on the test, whatever the test is, they need to make sure that it is benchmarked against the best in the world.

So that you are then flying with radar. You know how high that standard should be. If all the consortia do that, and if they are benchmarked against the best in the world, then they will be reasonably consistent and reasonably high, and that would close that loophole.

Senator FRANKEN. How do we guard against—how do we make them do that?

Mr. PHILLIPS. Well, first of all, I am assuming that when these standards are set, a lot of people are going to be watching. If the standards are set low, people like myself are going to write a lot of articles about it. Another thing you could do, you could build into the ESEA an evaluation of these activities, something like the National Academy of Sciences, where there could be an evaluation component where they would look at these things.

Senator FRANKEN. OK. I wanted to get to another thing, which is, I am very familiar with computer-adaptive tests, which I think are great, and they allow you to take them three times a year and study growth and actually use them as diagnostic tools so you can actually teach because of the results of tests. I have had principals refer to the No Child Left Behind tests that are taken in April, and you get the results back in June as "autopsies." So I understand the importance of those.

At the same time, Dr. Schmeiser talked about multiple assessments, and I am wondering how what you are talking about—the kind of tests you are talking about seem to be very objective, extremely objective, and can you do the other kind of multiple assessments with those, with the computer-adaptable tests, or does it mean that you have to take other tests, too? How do you reconcile these kind of two models, either of you?

Mr. PHILLIPS. I will start. I don't see that there is a problem because I can't imagine, except for some rare instances, tasks that could not be administered by computer. Some of them may not be scored by computer, but they could be administered by a computer, which cuts down on the cost and can also be adaptive.

I think there is a lot of flexibility and a lot of capacity and scalability with computer-adaptive testing that will make this consortia of States—a computer doesn't care whether you are testing a million students or 300 million students. It is scalable, and it makes this whole thing feasible. I don't see it as being incompatible with wanting to have multiple measures.

Senator FRANKEN. OK. Thank you.

I hope we do get another round, but if not, I will submit my other questions in writing. Thank you.

And thank you, Mr. Chairman.

The CHAIRMAN. Senator Isakson.

SENATOR ISAKSON

Senator ISAKSON. Thank you, Mr. Chairman.

Dr. Thurlow, what is your opinion of the current ESEA requirement on making AYP No. 1 and establishing a school or a system as "needs improvement?"

Ms. THURLOW. I can speak to that in terms of the tremendous benefits that we have seen for students with disabilities with a system that has set standards, held all students to those standards, and required that there be accountability for students. Perhaps one of the greatest advantages has been the requirement that we be able to see how subgroups are performing so that we can actually see how students with disabilities are doing.

I think that has been an advantage. That doesn't speak necessarily to opinions about AYP, etc. But the impetus behind that has been good.

Senator ISAKSON. Well, I know on page 6 of your testimony, the last sentence in the next to the last paragraph says, "Retrofitting assessments and accommodations and developing a series of alternative assessments because the general assessments do not work for all students is expensive for schools and stigmatizing for students."

I understand what that statement means. I am married to a special ed teacher. I have grown up—for 42 years. When you started your career, I was already married to a special ed teacher. I also know that there have been a lot kids, the 1 percent cognitive disability, which the chairman mentioned, I agree with your response about expanding that.

However, there are many different disabilities, and I have been advocating for a couple of years, really, when we get to this reauthorization, considering that the assessment of a special needs student be determined by the IEP that the parent and the special ed teachers develop, rather than being a singular assessment. What would you think about that?

Ms. THURLOW. Well, I would have concerns about that. The IEP has some very specific purposes related to laying out the goals for the student to get through individualized instructions. Those goals can be in many areas—behavioral supports, etc.

The IEP really is not a mechanism for accountability. It wasn't designed to do that. We would have different things all over the place, not just in different States, but in different districts and in different schools and in different classrooms based on IEP team members' understandings. I think it has been a tremendous advantage to have same standards for all students, and we would lose that.

Senator ISAKSON. Well, I am not talking about the—of course, I get that sometimes words mean different things to different people. I am not talking about the standards of the curriculum, but I am talking about the method of assessment of the achievement of the standards of the curriculum.

Ms. THURLOW. The IEP doesn't provide us with a method.

Senator ISAKSON. Is it not—wasn't it developed so the parent and the teacher got together to determine what was in the best interest of the child and their instruction for the coming year?

Ms. THURLOW. It is a legal document that helps parents work with educators to determine the specialized instruction that is

needed, what are the certain areas that we need to focus on, hopefully, to make sure that the student has access to the curriculum.

Senator ISAKSON. My reason for bringing this up is because in the last sentence in this paragraph that I read that you wrote, you could apply the same paragraph to "needs improvement" assessments on systems. Sometimes because of one disaggregated group, a system can become a "needs improvement" system or a school can become a "needs improvement" school. And I am a growth model guy. I think you ought to give schools a chance to work out of the stigma.

A lot of times special needs get the blame for that when, in fact, they are somewhat in gridlock because of the lack of any flexibility in what the assessment model will be. That is what I am getting at.

Ms. THURLOW. From my perspective, it is easy to blame a group when we don't know exactly how to make sure that they reach those standards, when we haven't figured out all of the ways to make sure that their achievement is improving.

Senator ISAKSON. Well, I would love to work with you on this subject because it is the single biggest thing that is going to affect IDEA and special needs as we come together, and I don't think this should become stigmatized, first of all. I don't think systems or schools should become stigmatized because we don't have the flexibility to assess the same standards for those kids so they get the same break, understanding the accommodation some of them are going to need.

Ms. THURLOW. Right. Just one last thing. We know that there are places where it is working, where students with disabilities are achieving, and the achievement gap is disappearing for students with disabilities. We have to look to those where they talk about shared understandings, collaboration, making sure they expect the same thing of all students, etc.

Thank you.

Senator ISAKSON. Thank you, Doctor.

Thank you, Mr. Chairman.

The CHAIRMAN. Senator Isakson, I look forward to working with you on this, too, because it is something that I know you care deeply about, and it is something that we have to focus on in the reauthorization of ESEA. I look forward to working with you on that.

Senator ISAKSON. Thank you.

The CHAIRMAN. Senator Bennet.

STATEMENT OF SENATOR BENNET

Senator BENNET. Thank you, Mr. Chairman. Thank you for this hearing.

Thank everybody for your testimony.

One of the, I think, unintended consequences of No Child Left Behind is that there is a horrible springtime ritual in this country, in our school districts, where we spend 2 or sometimes even 3 weeks administering what are largely standardized tests, having led up to that period of time with weeks and weeks and weeks, in some cases months of test preparation. Then that ritual ends, and throughout the rest of our 181-day school year, whatever it is, we do all the things we wanted to do during the school year. Like my

two daughters are practicing Shakespeare right now in their elementary school in Denver.

I think one of the causes—there are many causes of that. One of the causes of that is that we have had far too many standards at every grade level that have exhausted our kids and exhausted our teachers and not given us the information that we need, either for accountability purposes or for teaching and learning purposes.

I remember when I was superintendent in Denver, somebody came one day—I wish I could remember who it was, but I don't—and he had two books in his hand, and he was standing with our principal. He said, "This is your ninth grade math textbook. It is in English, and it costs $115. This is the ninth grade math textbook they use in Singapore, which parenthetically is in English and costs $15."

"The good news," he said, "is that this math textbook exists somewhere inside this math textbook, if you can only find it." Of course, what he was saying was we have to drill down in a much more rigorous way on fewer standards, fewer, clearer, higher, which is the purpose of your work, Dr. Paine, and the work of the other States.

I have two questions for anybody who wants to answer them. One, will we have accomplished finding this math textbook in these standards? Do we feel comfortable that we really are not covering the waterfront anymore, but we are going to do what is important, benchmark to an international norm? What can we do to help make sure that is true?

And second, what are the implications for technology or for test-making generally that may get us out of this springtime ritual that is so counterproductive for our kids?

Dr. Paine, maybe we will start with you?

Mr. PAINE. You sound like a school guy.

[Laughter.]

Very insightful. I went on a trip with some of our colleagues with CCSSO to Singapore, and the mantra is teach less, learn more. I think that the fewer, clearer, higher mantra in the——

Senator BENNET. Oh, I should say, Dr. Paine—sorry—that the point of that is that our bell regularly gets rung by the ninth grade kids in Singapore in terms of math results. So, sorry.

Mr. PAINE. Exactly. I think that reflective in the Common Core standards is the concept that, and I can speak from personal experience in West Virginia, that we have narrowed our State standards significantly, particularly in grades kindergarten through third grade, where we have reduced the numbers of standards, made them much more concise. We had the rigor, but we didn't have the simplicity, if you will, of concept so that we can really hone in on a few concepts well and lay that foundation for later grades.

That is one very practical thing that I think you will see in the Common Core set of standards. I do also agree with you with regard to this ritual that occurs every spring.

I have to tell you just a little story. I have just finished touring our State for about 3 weeks, conducting focus groups with the kids, teachers, parents, school superintendents, and local board presi-

dents. Interviewed each of our State board members, business community, PTA, and numerous, numerous people.

I got to the kids, and I talked to them about the State assessments and how much emphasis we place on the present model of using State assessments to assess our standards. I bring this up for a purpose so that I hope in the reauthorization we can find a way to fix this.

I was talking with the kids about this notion of linking teacher evaluation to student performance results and specifically the practice that seems to be emphasized right now is a summative test, which I really have to question, I have to be honest with you. I am all about accountability, but I think you have to be very careful about making those kinds of singular decisions on one particular assessment.

The kids said, "Well, that means on the State test, which we really don't take that seriously, we can take out Mr. Green." And I thought, "Oh, no." Then I said, "So you don't value that State test?" I said, "Do you value the NAEP?" I serve on the NAGBE board—governing board for the NAEP.

They said, "Well, sometimes we don't take that test as seriously." I won't tell you what they really said. So I said, "What test do you really value that will motivate you to learn all that you are taught daily?" And with respect to my colleague that sits to my left, they said, "The ACT." If we were an SAT State, they probably would have said that, too.

I think we have to figure out a way to merge purposes with a simpler testing strategy, if you will, in that springtime ritual that could be spread throughout the year and more frequent intervals as we look at assessing the Common Core so that we can make these assessments very important to our students. That is the point I wanted to bring out.

Mr. PHILLIPS. Well, let me address the technology question. The whole idea of technology is to make testing less burdensome and to get out of the way of instruction so you can have more instruction. I mean, just an obvious example of that, if you give a test that takes 2 months to get back to the students, you have made a lot of progress if you can get that result back in 15 seconds so the teacher can actually do something with it.

Technology is an important ingredient in this as we move to a new set of what I hope are State comparable assessments in order to make it feasible, in order to cut the costs down, and actually get better measurement because these type of tests—one of the reasons, for example, why there is a need—why there has been discussion for a 2 percent assessment is that existing paper/pencil tests give terrible measurements for that bottom set of students. So, there is a need to have a new test.

With a computer-adaptive test, it goes right down there and gets as good a measurement for them as it gets for everybody else. There is no need for a 2 percent assessment if you have a computer-adaptive assessment because it is doing as good a job for that bottom 2 percent as it is doing for the middle and for the top.

It is leveraging the technology that makes this feasible and practical and cost-effective in the future.

Senator BENNET. Thank you, Mr. Chairman.

The CHAIRMAN. Great questions and answers, provoked me to think about some questions.

Now let us see, Senator Hagan.

SENATOR HAGAN

Senator HAGAN. Thank you, Mr. Chairman.

I, too, want to thank you for having this hearing and certainly thank all of your witnesses for the time you spent preparing to come and sharing with us your thoughts and expertise.

I want to follow up on Senator Bennet's question. It wasn't my first question, but since we are right here. Dr. Phillips, what you just said on the computer-adaptive testing, that what it does is, it helps I guess from the bottom 2 percent just as much as the top 2 percent. Can you elaborate on what you mean by that?

Mr. PHILLIPS. Yes. A typical paper/pencil test, let us say, has got 40 items. Everybody takes the same 40 items. If you are a low-achieving student, that test is too hard for you. It doesn't do a good job of saying what you know. If you are a high-achieving student, it is too easy for you. It doesn't really measure—it sets a ceiling. You can't go any further.

What the computer-adaptive test does is it focuses on your level of ability, and it drills down and gets better and better and better measurements until it can't do any better. It therefore gives the same accuracy to a low-achieving student and to a high-achieving student.

Particularly, if you are measuring growth, which is one of the initiatives in the future, which has been mentioned, one of the things you see right now with paper and pencil tests, anytime you have growth measures, you always see the same phenomenon. High-achieving students do worse over time. Low-achieving students do better than you expect over time.

The reason for that is the high-achieving students are at the ceiling and can't go anywhere. The low-achieving students capitalize on chance, and they will bump up just due to chance. You get a much more accurate measurement of growth if you have the same precision and accuracy for high-achieving students as you have for low-achieving students. That is particularly important if you are going to hold teachers accountable.

If you are a teacher with a classroom of high-achieving students, it is going to be a disaster if you try to measure growth because they are likely to either not show any growth or show negative growth. Particularly when you get into the growth business, for low-achieving and high-achieving students, you need better measurement. That is what this would do.

Senator HAGAN. I think one of the things I have been concerned about is those high-achieving students, sometimes I don't think we expect as much out of them, that we have got to continue setting much, much higher expectations at the same time.

From your computer-adaptive testing, how many States are doing that right now, and what do you see to encourage other States to actually get onboard?

Mr. PHILLIPS. Many States around the country have some portion of their testing being done by computer adaptive. I could be wrong, but the three States I mentioned I think are the only ones that are

completely and totally computer adaptive. Oregon is the only State that has been actually approved through peer review and No Child Left Behind. The other two will go through that process shortly.

What I describe, shows that all testing companies are involved in this. There is a lot of innovation, R&D going on. If there was a signal from the Federal Government through ESEA or through the $350 million that this is important, there would be a whole lot more innovation and R&D, and the computer-adaptive tests in the future would be even better than they are today. Even today, they are practical and feasible and would give you what you need.

Senator HAGAN. One other comment that you said on that is that 15 minutes or whatever after the test is taken, then there could be some analysis in the States that use these.

Do the teachers actually then go back and do the students see these tests, see what they have done right and wrong? Because so many times, I think these students take these tests. You never see the booklets again. You never understand what you did wrong in order to evaluate it from that student's perspective.

Mr. PHILLIPS. Yes. What you could do with these tests, just to give you an example, in a typical paper and pencil test, let us say it is eighth grade and you are measuring the Pythagorean theorem, which is a subset of mathematics. You may only have two items that cover that. So you can't get a lot of good information. The teacher can't get a lot of good information to help determine if their students are learning the Pythagorean theorem.

With computer-adaptive testing, each student——

Senator HAGAN. I understand that, but do they actually go back and look at it?

Mr. PHILLIPS. Yes, they do.

Senator HAGAN. OK.

Mr. PHILLIPS. They get an immediate report, and you can see, "Oh, my students need to learn the Pythagorean theorem."

Senator HAGAN. I had one other question that I wanted to talk about just for a minute, and that is North Carolina is the first State in the Nation to create a Center for 21st Century Skills, with the goal of identifying those skills that will be most sought after in the workforce when—in the future with the idea to improve the States' education system to ensure that the students actually graduate with those skills.

Dr. Paine, I know in West Virginia, that you are also a leader in this effort, that West Virginia is. And I was just wondering, can you share with me any of what the Federal Government might do to encourage more States to identify and promote 21st century skills, and how can we sustain our State's commitment to this as new assessments continue to be put forward?

Mr. PAINE. Thank you.

Certainly, North Carolina was a leader, and we were the second State following North Carolina. I think that is a very insightful question. We tend to, if we are going to make decisions, we want to make sure we emphasize content, but embed higher-level skills—those "21st century" critical thinking, problem-solving, the IT skills, so forth—within the content.

Once again, when you do that, and that is what North Carolina has done as well, and that is what the business sector really wants

us to do in preparing a workforce. Whether it is a company with an international presence or whether it is a national company, a Fortune 500 company, or if it is a small business in West Virginia, I hear the same kinds of expectations. That calls for different teaching methodologies, which also calls for different assessment practices.

And in response to your previous question, because they both dovetail, there is one of the consortia that has developed is committed to the adaptive testing process, along with other types of measures. It is called Smarter Balance, and there are some 30 States that have come together to be a part of this consortia. It is being led by Sue Gendron, who is the former commissioner for Maine, who has just recently stepped down to head up this consortium.

Senator HAGAN. Thank you.

Thank you, Mr. Chairman.

The CHAIRMAN. Senator Merkley is gone.

Senator Casey.

STATEMENT OF SENATOR CASEY

Senator CASEY. Mr. Chairman, thank you and thanks for organizing this hearing and calling us together.

We want to thank our witnesses for your testimony and your work and your scholarship.

When I have traveled across Pennsylvania, as a State government-elected official and as a candidate and then in the time I have been in the Senate, whenever the topic of No Child Left Behind came up, it would usually be raised by others, and they would ask for a response. I found that people, whether they were mostly in the context of teachers and administrators, but others as well—I found that whenever the topic came up, people were even-tempered. They were always mad.

[Laughter.]

They had a real frustration, I think, with the gap between what was—what undergirded the original, the legislation itself and the expectations that flowed from that and then what the reality was when it was implemented. This is a broad kind of overview and too simplistic, but I will just frame it as simply as I can.

There was a sense that one side of the debate was yelling for more investment. The legislation made promises. The other side was saying we needed more measuring and assessment, and so that was implemented, or standards as well. It seems like we failed on both. We failed on the investment, and we failed on how we implemented the standards and assessments.

I know we don't have a lot of time, but I wanted to delve a little bit into the assessment question. Dr. Paine, I will start with you, and I invite others to comment as well. I found these two sentences among the many in your testimony on page 4 under the Common Assessment Development section.

You said, and I quote, in the first paragraph in that section, "Aligned standards and assessments will allow States aligned teacher preparation and other supports designed to improve overall student achievement and close achievement gaps." And then the sentence after that, which I thought was even more pointed in

terms of what we want to talk about. "Teachers from participating States will benefit from high-quality instructional supports and materials that are aligned to the core standards."

I wanted to focus your attention just on what are and how would you define and give examples of "high-quality instructional supports and materials?"

Mr. PAINE. I think you hit on perhaps the core of what could be the most important element, in my mind, of the reauthorization, and that is——

Senator CASEY. We didn't choreograph this either.

Mr. PAINE. No, sir. I have been very even-tempered about that, but——

[Laughter.]

I really think you hit on the issue, and that is developing the quality of teaching in our State, in our country is critical to the success of the education. We already know from research that that is probably the No. 1 variable that affects student achievement outside of what goes on outside of the home.

I think that we need to very quickly capitalize on a set of common standards and an assessment strategy to measure the full scope of that, which includes a variety of different measures, and help our teachers to understand how to become what I would call "assessment literate." Teach them how to read these standards and how to teach these standards and how to assess student achievement within those standards, and then to hold those accountable for their preparation.

And I think as we have one set of standards and hopefully can arrive at perhaps one set of assessments, perhaps two in two consortia, that allows us to really focus our efforts on how we do prepare teachers, and what are the best strategies for doing so? Then, how do we build a performance-based accountability system that makes some sense? We can do that with the collective energies of all the States.

Senator CASEY. Do you have any particular examples of those two words I mentioned, the materials and supports? Can you just amplify on that a little?

Mr. PAINE. That also allows a real focused effort on developing high-quality resources and materials in support of that Common Core. Instead of necessarily differences in the way that States, other resource partners, vendors, and so forth focus their efforts, all will be focused on that Common Core set of standards, which I think will capture a real positive synergy as we develop real rich, robust resources to support the teaching of those standards.

Senator CASEY. I know I am just about out of time. Anybody else want to quickly add to that?

[No response.]

Well, thanks very much.

Mr. Chairman, we will submit questions for the record as well. Thank you.

The CHAIRMAN. Thank you very much, Senator Casey.

Can I just return here to a couple of things? One specifically. Senator Bennet spoke about the annual spring exercise of testing and how much time it takes. Dr. Paine, if I am not mistaken, I

made some notes here, you said we need a simpler testing strategy, one that goes on during the school year.

One of the things that I have thought about for a long time is, how do teachers during the school year get an accurate assessment of each of their students that they are teaching on an ongoing basis so that you don't wait 2 or 3 or 4 months to find that something is happening here, and this student is not learning something?

It could be a simple thing like, well, in math it could be students are doing all right, but this one student, for some reason, is not doing very well. They seem to do OK in adding and multiplication and stuff, but they have a problem with fractions. If a teacher can find that out, then they can deal with this student and deal with the specificity of what it is that that student can't quite grasp.

I am familiar with a program that has been ongoing in Iowa that—at least I have heard from teachers who seem to love it—it has been kind of an experimental type program. It is technology-based, computer-based, where—and I played with it once a couple of years ago when it was just started to look at how if I were a teacher, how I would get this information.

It is a very rapid type of thing where literally on an almost daily basis or weekly basis, I should say, teachers get good information back about how their students are doing in each of these subjects. They also find out whether or not there is some part of that subject, maybe it is English, they are doing all right with punctuation. They seem to be doing all right with words and spelling, but they don't know where to place a verb. And they pick that up. The teacher gets that on an ongoing basis rather than just at a test at the end of the year that tests a more broad-based kind of achievement.

Are you familiar with any other kinds of programs; I am sure that is not the only one. Are there other kinds of programs that are technology-based, computer-based where you get this kind of simpler testing strategy that goes on during the year and doesn't just rely upon one or two big tests?

Mr. PAINE. There are programs and strategies that do exactly as you say and certain products that are produced by vendors out there in support of the teaching of standards. One of the projects that the Council of Chief State School Officers is undertaking, a very exciting initiative, I might add, is looking at what is the next generation of learners, and how do we support those needs?

Embedded within that concept, with the richness of technology that is now available to us, is to assess each student against each of these common standards and their progress in very real time so that teachers have access to that information via a very rich, robust data system on their desktop so that they can make those kinds of very frequent real-time decisions.

If you think about the possibilities of how that network then could be expanded to the home or to other places or a data system like that, we really have the capability to make those kinds of decisions. That is the undergirding of that kind of assessment system that I know you are referring to, those classroom assessments developed by teachers that are done in alignment with a broader assessment strategy that includes also a summative test.

The CHAIRMAN. Because one of the things that it seems to me that technology-based learning and the new technology, computer-based programs we have, kind of gets, to English language learners. Dr. Rivera, how has technology helped or hurt students who have to both be tested in English language learning, but also be tested in the core subjects that they have to learn also?

This is where I lack any knowledge. I don't really know whether or not technology has helped this. Has it assisted it? Have they focused on it? What is happening with new technologies in terms of English language learners?

Ms. RIVERA. I am not aware of efforts currently to develop assessments specifically. The English language proficiency tests, I think they are all given as paper and pencil, although perhaps there are some efforts to start developing them as computer-based assessments.

In terms of English language learners, what is going to be important in terms of the technology is to make sure that the schools that they are attending have access to the technology and instruction and that the instructional program integrates the technology and students are very capable of using the technology before we go off and try to test them using the technology.

I know that I worked a little bit on the standards at ACT actually on the writing assessment for NAEP, and the endeavor was to put NAEP on a computer-based platform. The committee I was on was to look at accommodations for English language learners and for students with disabilities. One of the cautions was to make sure that the instructional program really includes that kind of teaching. If it doesn't, then it is going to be problematic.

Also computers need to be available to students, and I know NAEP had a—perhaps it has been resolved. I know it was an issue in terms of the writing assessment that schools did not have the available computers to allow the testing to happen in an easy fashion.

The CHAIRMAN. Dr. Phillips, you indicated you wanted to address this?

Mr. PHILLIPS. Just in your earlier question, in the three States I mentioned, each of them have three opportunities for the student to take a test. In between those opportunities, the teacher can develop formative assessments also on a computer and get results on the same scale as a summative assessment.

So that if the student is having trouble with the Pythagorean theorem, she could say, "Well, what is it about the Pythagorean theorem you don't know?" and then develop an assessment based on that.

The other thing is your second question. In Oregon, their entire test for English language learners is computer-adaptive. It covers listening and speaking, and it is working just fine. There haven't been any issues.

Also, in terms of languages, there are no limit to the number of languages you can test in. The fact that we are doing it in English, that is just a random choice. You can do it in Spanish or any—for example, in Hawaii, we are doing it in Hawaiian.

The CHAIRMAN. Is it written, or is it spoken? Is it something that is an audible-type test?

Mr. PHILLIPS. In Oregon, it is both written and spoken.

The CHAIRMAN. Both.

Mr. PHILLIPS. Right. So you are assessing listening and speaking.

The CHAIRMAN. Interesting. Interesting.

Ms. RIVERA. It is an English language proficiency test that they have, right? Right. It is not a content assessment, which is different.

The CHAIRMAN. Right. Just for English language.

Ms. RIVERA. Right.

The CHAIRMAN. Senator Enzi.

Senator ENZI. Thank you, Mr. Chairman.

I will start with Dr. Rivera. For English language learners, are there accommodations that could be used for all types and forms of assessments that would maintain the validity of the scores for those students?

Ms. RIVERA. Well, that is an area we have been working with quite extensively, which is to identify what States are doing and what kinds of accommodations are available to States. There are accommodations, and we really have organized accommodations around the—for English language learners, the main thing that they need is access to the language of the test. They need to have linguistic access to the test. We call these linguistic accommodations.

States have many different kinds of accommodations, and they have policies that list accommodations. Often what they do, does not distinguish the accommodations for English language learners from students with disabilities. Making the decision as to which accommodation is appropriate for these students, I mean, Braille is not going to help an English language learner. Or moving things around, whatever. There are different kinds of accommodations.

It is very important that folks really have an understanding of what the needs are of the English language learner and that the appropriate accommodations are available to them, and those would be linguistic accommodations or accommodations that address the language, allow them access to the content of the test.

Senator ENZI. Thank you.

When Dr. Paine was speaking earlier, he reminded me of some student meetings that I have had. One of the things that really disturbed me is there is this general impression out there that there is no value in taking these tests. It doesn't matter how I do on it or what I do on it or even if I do it.

How do we overcome that? Did you come up with any great ideas based on your student/teacher work? And anybody else, too. If they take it seriously, it makes a difference in the scores, I suspect.

Mr. PAINE. Those students that are preparing to go to college certainly value college entrance tests. I think one of the secrets might be that we merge purposes of an assessment of the Common Core with a purpose, the same purpose or a shared purpose, excuse me, of college-going rates.

One other is that in our technical adult education classes, we are moving toward a 50 percent performance-based component, not just a paper/pencil test, as to whether you can be a good electrician. Now you will be juried by practitioners that will come in and actu-

ally assess your progress on a real, contextual, life-learning situation. Can you actually wire the house, so to speak?

If we can get at more performance-based types of assessments like that, those tend to engage kids, as you know. And move away from those traditional types of tests that kids are, quite frankly, tired of, that don't necessarily yield the kind of learning information that we need to know about kids in this 21st century. I think that is where this whole computer-adaptive and other strategies of assessment really provide tremendous hope as we assess the Common Core.

Senator ENZI. Thank you.

Anyone else?

Ms. RIVERA. I think that one of the very important pieces in all of this is the knowledge that the teachers have to be able to use the information that they gain from assessments. Also for them to be able to feel that they can develop classroom-based assessments and understand what skills their students need. That is really at the base, and it is very important.

It is not—teachers really don't feel comfortable often with assessment. Even if there is a rich body of data, they don't always feel comfortable being able to look at it and figure out what it is that it is really telling them about their students. We need to spend some time and some effort in helping teachers to understand how to use the information, how to use assessments appropriately.

I think in the new Race to the Top and the way these assessments are being developed, there is supposed to be an integrated system where there is summative assessment as well as perhaps benchmark assessments and classroom assessments. That whole system needs to be linked and connected, and teachers need to be able to have access to the data.

Senator ENZI. I have to tell you, all of that really bothers me. I thought that teachers were taught to assess and that that was their job in the classroom on a daily basis, and in that regard, they ought to be assessment literate. Why do we keep saying that the teachers don't know how to use the assessments?

The assessment may be bad. That still really bothers me. I will have some more questions that will deal with that.

Dr. Thurlow, quickly because my time has expired. Have students with disabilities been included in the development growth models, and if not, why not?

Ms. THURLOW. Growth models are complex, and I believe that students with disabilities have been included, if they participate in the regular assessment with accommodations that don't invalidate the results. So, yes, they have been included in that way. We have had students who are in the alternate assessment based on alternate achievement standards, and we haven't figured out how to include them very well. I think it is something that we are still working on.

Likewise, any other alternate assessment, unless it is based on the grade-level achievement standards, we haven't figured out very well how to include that in the growth model. Those are probably some of the students who most need to be included in a growth model.

Senator ENZI. Thank you.

The CHAIRMAN. Can I just add to that, those would be that 10 percent, or the 1 percent?

Ms. THURLOW. Yes. Yes.

The CHAIRMAN. Well, 1 percent, but it is 10 percent of students with disabilities.

Ms. THURLOW. Right.

The CHAIRMAN. Like a different slice, they would fit into what Senator Enzi was talking about?

Ms. THURLOW. Yes. Yes.

The CHAIRMAN. Senator Bennet.

Senator BENNET. Thank you, Mr. Chairman.

I just had two last questions. The first is that I think the State tests and NCLB has done a lousy job basically on accountability and a lousy job on teaching and learning because, among other things, we sort of push those two things together in the summative assessment that we are talking about, and I think it is really important for us to pull those apart.

Accountability is one piece of this puzzle, and it is different at this level of Government than it is at the State level and than it is at the district level. We are talking about measures of teacher effectiveness. Some of us are interested in differentiating pay. There are all kinds of things that fall into this category, and this category is teaching and learning and the ability of a teacher to assess her kids and then differentiate her instruction based on what she has seen to be able to meet the individual needs of the kids in her classroom.

Those are not the same thing, it seems to me. I wonder, as we are thinking about both the summative assessment at the end and the interim assessments or the benchmarks, the formative assessments, whatever it is we are talking about, whether we are giving thought to those distinctions? Are they important? Is this something we should be paying attention to from the schoolhouse level?

Does anybody have a reaction to that at all?

Mr. PAINE. I would, very quickly. You brought up an issue that I think is very, very important for us to address as we think about Common Core and how do we assess the Common Core? How do we support assessment literate teachers, and how do we then look at their performance relative to accountability measures?

I haven't met a teacher in our State that is not interested in making more money. It is how do I make more money? If you are going to link my performance to one single assessment, that could be problematic. We need to look at models that support a variety of student learning outcomes.

One of the issues that I would have with a typical—here we are calling them typical growth models now that assess annual progress, why don't growth models measure progress over more short, frequent intervals such as every 2 months where we drop several types of ways to assess progress at shorter, frequent intervals so we can inform kids, their parents, and teachers about their progress and incorporate all of that into some kind of a growth model of accountability?

Senator BENNET. Is there anybody else?

Ms. SCHMEISER. I would add a quick comment that I agree totally with the last comment. I would also say that the Common

Core standards allow us an opportunity to align our systems in this country like we have never been able to do before, grade by grade, in an aligned, coherent way, looking at student growth longitudinally over time.

Inside of that can be multiple measures that can be embedded in instruction. They can be benchmark examinations. They can be summative. When you have a common goal and when it is clear what kids need to know and be able to do at the end of the third grade, fourth grade, fifth grade, you open doors for being able to have an aligned system both every day in the classroom to improve instruction all the way through the system.

I would say, Senator Bennet, I think it is very important to begin to think about the roles that assessment can play in the classroom, as well as for different purposes at different times, and make sure from the very beginning they are well-planned and well-aligned as we look at longitudinal student progress.

Senator BENNET. Well, what is interesting about that is that the more aligned, if you imagine a system that is perfectly aligned—I don't think there is one. I have never seen it in the United States. It ought to be a system, sort of ironically, as a consequence of that alignment that allows the system to differentiate to the maximum degree. That is really what we are talking about.

Ms. SCHMEISER. Yes, that is right.

Senator BENNET. The last question I had was for Dr. Phillips because I am least familiar with the things that you have talked about today, and I appreciate learning about it.

Has anybody done an analysis of the capacity of the school districts in this country from a technological point of view to administer what you are talking about? Because I suspect there is huge variability in the United States of how many computers are available, the wiring, and all of that. I was just curious whether there is something that I could look at and read about that?

Mr. PHILLIPS. There are surveys. And what was found in the three States I mentioned is that that is not really an issue because since the testing window is the whole year, you don't have to have a computer for every student at the same time. And in the rare cases where a district or a school doesn't have a computer, this is obvious leverage to get them one. It is a kind of a win-win situation.

Even in Hawaii, where even the most remote islands, I think we only found maybe one case where they needed a computer and didn't have one. In the old days when you thought about computer-adaptive testing, this was an issue, but it is really not an issue when the testing window is all year long.

Ms. THURLOW. I would refer you also to a study that NAEP did. I think in 2008, they collected information about technology, and it was quite positive. I can't remember all the facts, but I could certainly get the reference to you.

Senator BENNET. Thank you, Mr. Chairman.

And thank you to the witnesses.

The CHAIRMAN. Thank you, Senator Bennet.

Senator ENZI. I have questions, but I will submit them.

The CHAIRMAN. Well, I think this has been a very informative panel. I join with Senator Bennet and Senator Enzi and all the rest

of the Senators in thanking you for your excellent testimony, both written and verbal.

We will leave the record open for 10 days for other questions that we might want to submit to you in writing, appreciate if you would answer those.

I also ask that you continue to keep us informed as we move along in ESEA reauthorization, your suggestions, your advice. I am sure that through your different networks, you will know what we are doing here. I hope that you will continue to inform and advise us as we move along.

And I hope that we, in turn, our staffs can continue to be in touch with you as we develop this.

So thank you very much. Great hearing. Appreciate it.

The committee will stand adjourned.

[Editor's Note: The following report was requested to be included in the hearing record: Policy Analysis—Behind the Curtain: Assessing the Case for National Curriculum Standards by Neal McCluskey. This report may be found at *http://www.cato.org/pub_display.php?pub_id=11217*.

Due to the high cost of printing, materials that have been previously published are not reprinted in the hearing record.]

[Whereupon, at 4:02 p.m., the hearing was adjourned.]

○