

**Testimony before the House Education and Labor Committee on
the Re-Authorization of No Child Left Behind**

Linda Darling-Hammond

September 10, 2007

Congressman Miller, Congressman McKeon and members of the Committee. Thank you for this opportunity to testify on the draft bill to re-authorize No Child Left Behind. I am Linda Darling-Hammond, Charles E. Ducommun Professor of Education at Stanford University and co-director of the Stanford Educational Leadership Institute and the School Redesign Network. I was also the founding Executive Director of the National Commission on Teaching and America's Future, and have spent many years studying policies and practices in the U.S. and around the world that support stronger curriculum, assessment, teaching and learning.

I want also to thank the Committee for its openness and commitment to the democratic process in having shared a public draft of the re-authorization bill prior to finalizing the bill. This move shows a respect and consideration for the public that is appreciated by those who care deeply about our nation's education system.

While the very complex NCLB legislation has many elements that deserve attention and ongoing revision, I am sure you will hear about those from many others. I want to focus my testimony this morning on three key elements of the law:

1. The provisions to encourage multiple measures of assessment and multiple indicators of school progress, which I believe are essential to raise standards and strengthen educational quality in ways that are internationally competitive;
2. The provisions to improve the quality and distribution of the teaching force, which are also essential to our ability to reach the high goals this Congress would like to establish for our nation's schools, and
3. The means for measuring school progress from year to year, which I believe need to become more publicly comprehensible and more closely focused on evaluating continuing progress for students and schools.

My comments are based on studies of U.S. education and of the education systems of other countries that are outperforming the U.S. by larger and larger margins every year. For example, in the most recent PISA assessments, the U.S. ranked 19th out of 40 countries in reading, 20th in science, and 28th in math (on a par with Latvia), outscored by nations like Finland, Sweden, Canada, Hong Kong, South Korea, the Netherlands, Japan, and Singapore (which did not participate in PISA but scored at the top of the TIMSS rankings) that are investing intensively in the kinds of curriculum and assessments and the kinds of teaching force improvements that we desperately need and that this re-authorization bill is seeking to introduce.

2003 PISA RESULTS

<u>Reading</u>	<u>Scientific Literacy</u>	<u>Math</u>
Finland	Finland	Hong Kong
South Korea	Japan	Finland
Canada	Hong Kong	South Korea
Australia	South Korea	Netherlands
Liechtenstein	Liechtenstein	Liechtenstein
New Zealand	Australia	Japan
Ireland	Macao	Canada
Sweden	Netherlands	Belgium
Netherlands	Czech Republic	Macao (China)
U.S. ranks # 19 / 40	U.S. ranks #20 / 40	U.S. ranks #28 / 40

It is worth noting that PISA assessments focus explicitly on 21st century skills, going beyond the question posed by most U.S. standardized tests, “Did students learn what we taught them?” to ask, “What can students do with what they have learned?” PISA defines literacy in mathematics, science, and reading as students’ abilities to *apply* what they know to new problems and situations. This is the kind of higher-order learning that is increasingly emphasized in other nations’ assessment systems, but often discouraged by the multiple-choice tests most states have adopted under the first authorization of No Child Left Behind. Underneath the United States’ poor standing is an outcome of both enormous inequality in school inputs and outcomes and a lack of sufficient focus for all students on higher-order thinking and problem-solving, the areas where all groups in the U.S. do least well on international tests.

In addition to declines in performance on international assessments, the U.S. has slipped in relation to other countries in terms of graduation rates and college-going. Most European and Asian countries that once educated fewer of their citizens now routinely graduate virtually all of their students. Meanwhile, the U.S. has not improved graduation rates for a quarter century, and graduation rates are now going *down* as requirements for an educated workforce are going steeply up. According to an ETS study, only about 69% of high school students graduated with a standard diploma in 2000, down from 77% in 1969 (Barton, 2005). Of the 60% of graduates who go onto college, only about half graduate from college with a degree. In the end, less than 30% of an age cohort in the U.S. gains a college degree (U.S. Census Bureau, 2005). For students of color, the pipeline leaks more profusely at every juncture. Only about 17% of African American young people between the ages of 25 and 29 – and only 11% of Hispanic youth -- had earned a college degree in 2005, as compared to 34 % of white youth in the same age bracket (U.S. Census Bureau, 2005).

And whereas the U.S. was an unchallenged 1st in the world in higher education participation for many decades, it has slipped to 13th and college participation for our young people is declining (Douglass, 2006). Just over one-third of U.S. young adults are participating in higher education, most in community colleges. Meanwhile, the countries belonging to the Organization for Economic Cooperation and Development (OECD), which are mostly European, now average nearly 50% participation in higher education,

and most of these students are in programs leading to a bachelors degree. Similarly in Southeast Asia, enormous investments in both K-12 and higher education have steeply raised graduation rates from high school and college-going rates.

The implications of these trends are important for national economies. A recent OECD report found that for every year that the average schooling level of the population is raised, there is a corresponding increase of 3.7% in long-term economic growth (OECD, 2005), a statistic worth particular note while the U.S. is going backwards in educating its citizens, and most of the rest of the world is moving forward.

What are High-Achieving Nations Doing?

Funding. Most high-achieving countries not only provide high-quality universal preschool and health care for children, they also fund their schools centrally and equally, with additional funds to the neediest schools. By contrast, in the U.S., the wealthiest school districts spend nearly ten times more than the poorest, and spending ratios of 3 to 1 are common within states (ETS, 1991; Kozol, 2005). These disparities reinforce the wide inequalities in income among families, with the most resources being spent on children from the wealthiest communities and the fewest on the children of the poor, especially in high-minority communities.

Teaching. Furthermore, high-achieving nations intensively support a better-prepared teaching force – funding competitive salaries and high-quality teacher education, mentoring, and ongoing professional development for all teachers, at government expense. Countries which rarely experience teacher shortages (such as Finland, Sweden, Norway, Netherlands, Germany, Australia, New Zealand, Japan, Taiwan, Singapore) have made substantial investments in teacher training and equitable teacher distribution in the last two decades. These include:

- High-quality pre-service teacher education, completely free of charge to all candidates, including a year of practice teaching in a clinical school connected to the university,
- Mentoring for all beginners in their first year of teaching from expert teachers, coupled with other supports like a reduced teaching load and shared planning,
- Salaries which are competitive with other professions, such as engineering and are equitable across schools (often with additional stipends for hard-to-staff locations),
- Ongoing professional learning embedded in 10 or more hours a week of planning and professional development time (Darling-Hammond, 2005).

Leaders in Finland attribute the country's dramatic climb from the bottom of the international rankings to the very top to intensive investments in teacher education. Over ten years the country overhauled preparation to focus more on teaching for higher-order skills and teaching diverse learners – including a strong emphasis on those with special needs – and created a funding stream to provide a 3-year graduate level preparation program to all teacher candidates free of charge and with a living stipend, a full year of training in a professional development school site – rather like the residency promoted in this draft bill, intensive mentoring once in the classroom, and more than ten hours a week of professional learning time in school, where teachers collaborate on lesson planning and on the development and scoring of local performance assessments that are the backbone of the country's assessment system.

In high-achieving Singapore, which I recently visited as part of a review team for the Institute of Education, students from the top 1/3 of the high school class are recruited into a 4-year teacher education program (or, if they enter later, a one-year graduate program) and immediately put on the Ministry's payroll as employees. They are paid a stipend while they are in training (which is free for them) and are paid at a rate that is higher than beginning doctors when they enter the profession. There they receive systematic mentoring from expert teachers once they begin teaching. Like all other teachers in Singapore, the government pays for 100 hours of professional development annually in addition to the 20 hours a week they have to work with other teachers and visit each others' classrooms to study teaching. As they progress through the career, there are 3 separate career ladders they can pursue, with support from the government for further training: developing the skills and taking on the responsibilities of curriculum specialists, teaching / mentoring specialists, or prospective principals.

Curriculum and Assessment. Finally, these high-achieving nations focus their curriculum on critical thinking and problem solving, using examinations that require students to conduct research and scientific investigations, solve complex real-world problems in mathematics, and defend their ideas orally and in writing. In most cases, their assessment systems combine centralized (state or national) assessments that use mostly open-ended and essay questions and local assessments given by teachers, which

are factored into the final examination scores. These local assessments – which include research papers, applied science experiments, presentations of various kinds, and projects and products that students construct -- are mapped to the syllabus and the standards for the subject and are selected because they represent critical skills, topics, and concepts. They are often suggested and outlined in the curriculum, but they are generally designed, administered, and scored locally.

An example of such assessments can be found in Appendix A, which shows science assessments from high-achieving Victoria, Australia and Hong Kong – which use very similar assessment systems -- in comparison to traditional multiple choice or short answer items from the United States. Whereas students in most parts of the U.S. are typically asked simply to memorize facts which they need to recognize in a list answers, or give short answers which are also just one-sentence accounts of memorized facts, students in Australia and Hong Kong (as well as other high-achieving nations) are asked to apply their knowledge in the ways that scientists do.

The item from the Victoria, Australia biology test, for example, describes a particular virus to students, asks them to design a drug to kill the virus and explain how the drug operates (complete with diagrams), and then to design an experiment to test the drug. This state test in Victoria comprises no more than 50% of the total examination score. The remaining components of the examination score come from required assignments and assessments students undertake throughout the year – lab experiments and investigations as well as research papers and presentations – which are designed in response to the syllabus. These ensure that they are getting the kind of learning opportunities which prepare them for the assessments they will later take, that they are getting feedback they need to improve, and that they will be prepared to succeed not only on these very challenging tests but in college and in life, where they will have to apply knowledge in these ways.

Locally managed performance assessments that get students to apply their knowledge to real-world problems are critically important to the teaching and learning process. They allow the testing of more complex skills that cannot be measured in a two-hour test on a single day. They shape the curriculum in ways that ensure stronger learning opportunities. They give teachers timely, formative information they need to

help students improve -- something that standardized examinations with long lapses between administration and results cannot do. And they help teachers become more knowledgeable about the standards and how to teach to them, as well as about their own students and how they learn. The process of using these assessments improves their teaching and their students' learning. The processes of collective scoring and moderation that many nations or states use to ensure reliability in scoring also prove educative for teachers, who learn to calibrate their sense of the standards to common benchmarks.

The power of such assessments for teaching and learning is suggested by the fact that ambitious nations are consciously increasing the use of school-based performance assessments in their systems. Hong Kong, Singapore, and several Australian states have intensive efforts underway to expand these assessments. England, Canada, Sweden, and the Netherlands have already done so. Locally managed performance assessments comprise the entire assessment system in top-ranked Finland and in Queensland and ACT, Australia – the highest-achieving states in that high-achieving nation.

These assessments are not used to rank or punish schools, or to deny promotion or diplomas to students. (In fact, several countries have explicit proscriptions against such practices). They are used to evaluate curriculum and guide investments in professional learning -- in short, to help schools improve. By asking students to show what they know through real-world applications of knowledge, these other nations' assessment systems encourage serious intellectual activities that are currently being discouraged in U.S. schools by the tests many states have adopted under NCLB.

How NCLB can Help the United States Become Educationally Competitive

Multiple Measures and Performance Assessments. The proposals in the re-authorization draft to permit states to use a broader set of assessments and to encourage the development and use of performance assessments are critical to creating a globally competitive curriculum in U.S. schools. We need to encourage our states to evaluate the higher-order thinking and performance skills that leading nations emphasize in their systems, and we need to create incentives that value keeping students in school through graduation as much as producing apparently high average scores at the school level.

Many states developed systems that include state and locally-administered performance assessments as part of their efforts to develop standards under Goals 2000 in the 1990s. (These states included Connecticut, Kentucky, Maine, Maryland, Nebraska, New Hampshire, New Jersey, New York, Oregon, Vermont, Rhode Island, Washington, Wisconsin, and Wyoming, among others.) Not coincidentally, these include most of the highest-achieving states in the U.S. on the National Assessment of Educational Progress. Indeed, the National Science Foundation provided millions of dollars for states to develop such hands-on science and math assessments as part of its Systemic Science Initiative in the 1990s, and prototypes exist all over the country. One such measure -- a science investigation requiring students to design, conduct, analyze, and write up results for an experiment -- currently used as a state science assessment in Connecticut (a top-ranked state in both science and writing) is included with the assessment examples in Appendix A.

Researchers learned that such assessments can be managed productively and reliably scored with appropriate training and professional development for teachers, along with moderation and auditing systems, and that teaching and student achievement improve when such assessments are used. (For a review, see Darling-Hammond & Rustique-Forrester, 2006).

However, the initial years of NCLB have discouraged the use and further development of these assessments, and have narrowed the curriculum both in terms of the subjects and kinds of skills taught. NCLB's rapidly implemented requirement for every-child every-year testing created large costs and administrative challenges that have caused some states to abandon their performance assessments for machine-scored, multiple choice tests that are less expensive to score and more easily satisfy the law. In addition, the Department of Education has discouraged states from using such assessments. When Connecticut sued the federal government for the funds needed to maintain its sophisticated performance assessments on an every-child every-year basis, the Department suggested the state drop these tasks -- which resemble those used in high-scoring nations around the world -- for multiple choice tests. Thus the administration of the law is driving the U.S. curriculum in the opposite direction from what a 21st century economy requires.

A number of studies have found that an exclusive emphasis on (primarily multiple-choice) standardized test scores has narrowed the curriculum. The most recent reports of the Center for Education Policy (CEP) and the National Center for Education Statistics (May 2007 *Stats in Brief*) confirm sizeable drops in time dedicated to areas other than reading and math, including science, history, art, and physical education. The CEP also found that districts are more tightly aligning their instruction to this limited format as well as content of state tests. While these tests are one useful indicator of achievement, studies document that they often overemphasize low-level learning. As reporter Thomas Toch recently stated, "The problem is that these dumbed-down tests encourage teachers to make the same low-level skills the priority in their classrooms, at the expense of the higher standards that the federal law has sought to promote." To succeed in college, employment and life in general, students need critical thinking and problem solving skills that the tests fail to measure, and they need a complete curriculum.

Teachers in many states report that the curriculum is distorted by tests and that they feel pressured to use test formats in their instruction and to teach in ways that contradict their ideas of sound instructional practice. An *Education Week* survey of more than 1,000 public school teachers reported that two-thirds felt their states had become too focused on state tests; 85% reported that their school gives less attention to subjects that are not on the state test. One Texas teacher noted, "At our school, third- and fourth-grade teachers are told not to teach social studies and science until March." Teachers often feel that their responses to tests are not educationally appropriate. These comments from teachers – reflecting the view of a majority in recent surveys -- give a sense of the problem:

Before [our current state test] I was a better teacher. I was exposing my children to a wide range of science and social studies experiences. I taught using themes that really immersed the children into learning about a topic using their reading, writing, math, and technology skills. Now I'm basically afraid to NOT teach to the test. I know that the way I was teaching was building a better foundation for my kids as well as a love of learning. Now each year I can't wait until March is over so I can spend the last two and a half months of school teaching the way I want to teach, the way I know students will be excited about.

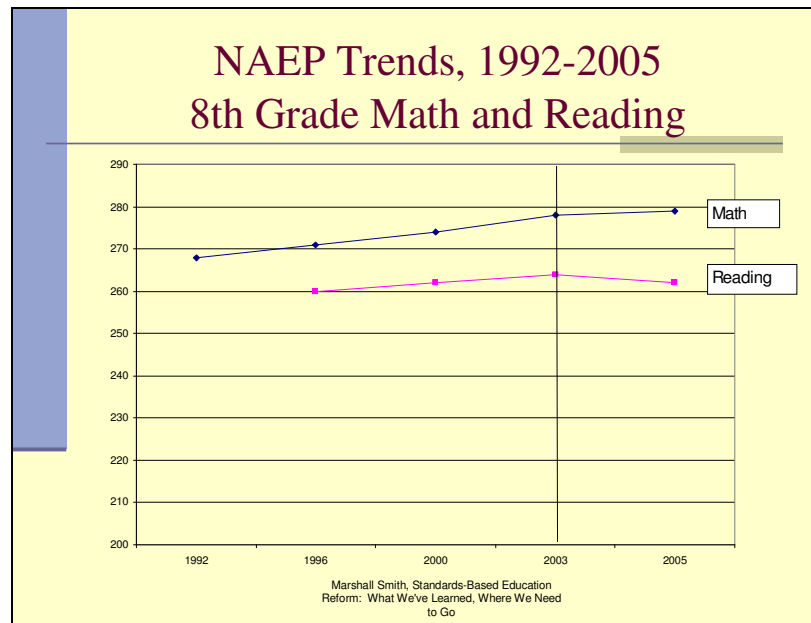
■ A Florida Teacher

I have seen more students who can pass the [state test] but cannot apply those skills to anything if it's not in the test format. I have students who can do the test but can't look up words in a dictionary and understand the different meanings.... As for higher quality teaching, I'm not sure I would call it that. Because of the pressure for passing scores, more and more time is spent practicing the test and putting everything in test format.

■ A Texas Teacher

Studies find that, as a result of test score pressures, students are doing less extended writing, science inquiry, research in social sciences and other fields, and intensive projects that require planning, finding, analyzing, integrating, and presenting information – the skills increasingly needed in a 21st century workforce. The use of computers for writing and other purposes has even declined in states that do not allow computer use on their standardized tests (for a summary, see Darling-Hammond & Rustique-Forrester, 2005). This narrowing is thought to be one reason for the poor performance of the U.S. on international assessments like PISA, which evaluate how students can apply knowledge to complex problems in new situations.

Indeed, as state test scores have gone up under NCLB, scores on other tests measuring broader skills have not. For example, for some states, Reading gains are positive on the state test but negative on the more intellectually challenging NAEP test. Overall, data from the trend NAEP assessment show that math gains from the 1990s have leveled off since 2002 and reading has declined.



Perhaps the most troubling unintended consequence of NCLB has been that the law creates incentives for schools to boost scores by pushing low-scoring students out of school. The very important goal of graduating more of our students has simply not been implemented, and the accountability provisions actually reward schools with high dropout rates. Push-out incentives and the narrowed curriculum are especially severe for students with disabilities, English language learners, students of color and economically disadvantaged students. Recent reports of the Public Education Network confirm that parents, students and other community members are concerned about the over-reliance on test scores for evaluating students and schools. A number of recent studies have confirmed that this over-reliance has been associated with grade retention and other school actions that exacerbate dropout rates and student exclusion from school, especially for low-income students of color.¹ This creates the perverse outcome that efforts to raise standards are resulting in fewer students receiving an education.

If education is to improve in the United States, schools must be assessed in ways that produce high-quality learning and that create incentives to keep students in school. A central part of a solution to these problems is to employ multiple forms of assessment and multiple indicators, while retaining the powerful tools of publicly available assessment information and the critically important focus on equity. The provisions of the draft bill that allow states to develop and use such measures, and the requirements that these include graduation rates, are essential to creating the incentives for a world-class curriculum within a world-class education system that actually reduces the achievement gap while ensuring more and more students are well-educated. A multiple measures approach can help schools and districts improve student outcomes more effectively because:

1. The use of multiple measures ensures that attention will be given to a comprehensive academic program and a more complete array of important learning outcomes;
2. A multiple measures approach can incorporate assessments that evaluate the full range of standards, including those addressing higher-order thinking and performance skills;
3. Multiple measures provide accountability checks and balances so that emphasizing one measure does not come at the expense of others (e.g. boosting

test scores by excluding students from school), but they can give greater emphasis to priority areas; and

4. A multiple measures index can provide schools and districts with incentives to attend to the progress of students at every point on the achievement spectrum, including those who initially score far below or above the test score cut point labeled “proficient.” It can encourage schools to focus on the needs of low-scoring students, students with disabilities, and ELL students, using assessments that measure gains from wherever students begin and helping them achieve growth.

One of the central concepts of NCLB’s approach is that schools and systems will organize their efforts around the measures for which they are held accountable. Because focusing exclusively on a single indicator is both partial and problematic, the concept of multiple measures is routinely used by policymakers to make critical decisions about such matters as employment and economic forecasting (e.g., the Dow Jones Index or the GNP), as well as admissions to college. Successful businesses use a “dashboard” set of indicators to evaluate their health and progress, aware that no single measure is sufficient to understand or guide their operations. Business leaders understand that efforts to maximize short-term profits alone could lead to behaviors that undermine the long-term health of the enterprise.

Similarly, use of a single measure to guide education can create unintended negative consequences or fail to focus schools on doing those things that can improve their long-term health and the education of their students. Indeed, the measurement community’s *Standards for Educational and Psychological Testing* mandates the use of multiple sources of evidence for major decisions. NCLB calls for multiple measures of student performance, and some states have developed systems that incorporate such measures. Up to now, implementation of the law has not promoted their use for evaluating school progress. In the new NCLB these and other states will be supported to develop systems that resemble those in the highest-achieving nations around the globe.

Multiple indicators can counter the problems caused by over-reliance on single measures. Multiple forms of assessment can include traditional statewide tests as well as other assessments, developed at the state or local levels, that include writing samples, research projects, and science investigations, as well as collections of student work over time. These can be scored reliably according to common standards and can inform

instruction in order to improve teaching and learning. Such assessments would only be used for accountability purposes when they meet the appropriate technical criteria, reflect state-approved standards, and apply equitably to all students, as is already the case in Connecticut, Nebraska, Oregon, Vermont, and other states successfully using multiple forms of assessment.

To counter the narrowing of the curriculum and exclusion of important subjects that has been extensively documented as a consequence of NCLB, the new law should, as this draft proposes, allow states to include other subjects, using multiple forms of assessment, in an **index of school indicators** measuring school progress toward a “proficiency benchmark” that incorporates both good measures of learning and measures of graduation and progress through school. To ensure strong attention is given to reading and math, these subjects can be weighted more heavily. Graduation rates and grade promotion rates should be given substantial weight in any accountability system. Other relevant indicators of school progress, such as attendance or participation in rigorous coursework, could be included. (For specifics on how such an index might operate, see Appendix B.) An index that tracks and sets targets for continual school progress – including the progress of student groups within the school – at all points along the achievement continuum would accomplish several goals:

- It would actually measure how much students are learning, taking into account the progress of all students not just a select few, including students who score well above or below the “proficiency” level;
- It would allow for more appropriate attention to and assessment of special education students and English language learners;
- It would provide incentives for schools to offer a full curriculum and to incorporate multiple measures of learning that include more ambitious performance assessments;
- It would provide a better warning system, distinguishing between schools that are making steady progress and those that are truly failing and thus unable to make progress on the index, so that states can focus on those needing the most help
- It would enable teachers and schools to chart students’ progress and increase ambitions for all, to proficiency and beyond

- It would create incentives for schools to invest in all students' education, to keep students in school, and to address all aspects of performance.

Because evidence is clear that multiple assessments are beneficial to student learning and accountability decisions, it is promising that the bill includes a provision to provide significant funds to assist states and districts to implement systems that include multiple forms of evidence about student learning, including state and local performance assessments. of state assessment and accountability systems.

These points in support of a multiple measures approach to accountability were made in two recent letters to the Congress -- one from a group of 23 leading civil rights organizations, including Aspira, LULAC, the NAACP, the National Council for Educating Black Children, and others, and the other from more than 120 leading educational experts, including the nation's most prominent testing experts and more than a dozen former presidents of organizations including the American Educational Research Association, the National Academy of Education, and the National Council for Measurement in Education. These letters can be found <http://www.edaccountability.org>.

Investments in Teaching. Once we develop a strong curriculum that focuses on 21st century skills, which teaches and assesses the skills we need in the ways that students will use them in the real world, we must also ensure that we have well-prepared and well-supported teachers who know and can teach challenging content extremely well to the very diverse group of students in our schools. Few of the conditions that support teaching in high-achieving nations are routinely in place in school systems across the U.S. and they are especially lacking in the school districts and schools which serve most low-income students and students of color.

Unfortunately, unlike other industrialized nations that are high-achieving, the United States lacks a systematic approach to recruiting, preparing, and retaining teachers, or for using the skills of accomplished teachers to help improve schools. With unequal resources across districts, and few governmental supports for preparation or mentoring, teachers in the U.S. enter:

- with dramatically different levels of training -- with those least prepared typically teaching the most educationally vulnerable children,

- at sharply disparate salaries -- with those teaching the neediest students typically earning the least,
- working under radically different teaching conditions – with those in the most affluent communities benefiting from class sizes under 20 and a cornucopia of materials, equipment, specialists, and supports, while those in the poorest communities teach classes of 40 or more without adequate books and supplies,
- with little or no on-the-job mentoring or coaching in most communities to help teachers improve their skills

Most also have few ways to engage in developing and using their skills to maximum advantage, spending most of their careers teaching solo in egg-crate classrooms, rather than working with colleagues to improve curriculum, instruction, and assessment.

This re-authorization proposal promises to make substantial headway on these problems. Particularly important are several elements of the TEACH Act that have been integrated into the bill. These include:

- Recruitment incentives to attract both well-prepared novices and accomplished veteran teachers into high-need schools, through innovative training and compensation approaches;
- Improvements in teachers' preparation through new teacher residency programs in high need communities, as well as improvements in all teachers' preparation to teach content standards, to teach diverse students well, and to use technology;
- A focus on improving teacher education and teacher effectiveness through the development of a nationally available teacher performance assessment
- High-quality mentoring for all beginning teachers;
- Strong professional development through new Teacher and Principal Professional Development Academies; and
- The development of career ladders for teachers that can recognize and reward highly-accomplished and effective teachers who show high levels of performance and the ability to contribute to student learning – and that can take advantage of these teachers' expertise by creating mentor and master teacher positions that allow them to support other teachers and the school as a whole in improving curriculum and instruction.

This comprehensive approach can begin to transform our conceptions of the teaching career in much the way that other countries have already done system-wide.

Many elements of the bill are based on a thoughtful diagnosis of our teacher supply and quality problems and a set of initial steps that, if eventually integrated system-wide, could actually begin to solve these problems. Below I touch briefly on the reasons for the importance and likely success of these elements of the bill.

Recruitment Incentives to Attract Expert Teachers to High-Need Schools –

Much research has shown that teachers are the most unequally distributed school resource and that low-income schools have a disproportionate number of inexperienced and under-prepared teachers. Recruitment incentives for high-need schools are needed to attract and keep expert, experienced teachers in the schools where they are most needed, both to teach and to mentor other teachers. The bill offers a combination of salary incentives and improvements in working conditions, including time for teachers to work and plan together, which have been shown to influence teachers' career decisions.

As part of a broader career ladder initiative, federal matching grants to states and districts can provide incentives for the design of innovative approaches to attract and keep accomplished teachers in priority low-income schools, through compensation for accomplishment and for additional responsibilities, such as mentoring and coaching. The bill allows for districts to recognize teacher expertise through such mechanisms as National Board Certification, state or local standards-based evaluations, and carefully assembled evidence of contributions to student learning.

Improvements in Teacher Preparation and Professional Development. While NCLB's highly qualified teacher provision has strengthened preparation in the content areas, there is much work to be done to improve teacher effectiveness. Major needs are stronger preparation for teachers to learn how to teach effectively within their content areas, how to design and use assessments that reveal how students are learning and guide teaching, how to teach reading and literacy skills at all grade levels, and how to teach special education students and English language learners. These students are the disproportionately ones who are failing to meet standards under NCLB and their teachers need very sophisticated skills to help them.

The TEACH Act proposes grants to strengthen teacher preparation and professional development in these areas which represent best practices in the field – involving teachers in curriculum and assessment planning, modeling and demonstration

of teaching strategies, and follow up coaching in classrooms in both pre- and in-service development programs. These approaches should replace the “hit-and-run” professional development that is currently common. **Professional Development Academies** can provide a steady supply of high-quality professional development of the kind that has been shown to improve student achievement – intensive institutes and study opportunities for networks of teachers who can both work on these practices together and receive on the job coaching to hone their skills.

New models of teacher preparation are especially needed in our high-need districts. The most critical need for improving teacher preparation is to ensure that programs provide one of the most important elements of preparation – the opportunity to learn under the direct supervision of expert teachers working in schools that serve high-need students well. Teaching cannot be learned from books or even from being mentored periodically. Teachers must see expert practices modeled and must practice them with help. However, student teaching is too often reduced or omitted, or it is in classrooms that do not model expert practice, or it is in classrooms that do not serve high-need students – and what is learned does not generalize to other schools. This fundamental problem has to be tackled and solved if we are to prepare an adequate supply of teachers who will enter urban or poor rural classrooms competent to work effectively with the neediest students and confident enough to stay in teaching in these areas.

The Bill provides for **teacher residency programs** in high-need communities. This alternative has proven successful in the Urban Teacher Residency designed in Chicago that has created new schools or completely re-staffed existing schools with highly expert mentor teachers (like professional development schools) and then placed mid-career recruits *in the classrooms of* these mentor teachers for a year while they complete coursework in curriculum, teaching, and learning at partner universities. Rather than trying to teach without seeing good teaching in a sink or swim model, these recruits watch experts in action and are tutored into accomplished practice. They receive a \$30,000 salary during this year and a master’s degree and credential at the end of the year. They continue to receive mentoring in the next two years. They are selected because they want to commit to a career in urban public school teaching and they pledge to spend at least four years in city schools. This model has already shown high retention rates in

teaching and strong performance by graduates, who now staff other turnaround schools in the city. Similar models have been launched in Boston and other cities. Such programs can solve several problems simultaneously – creating a pipeline of committed teachers who are well-prepared to engage in best practice for children in for high-need schools, while creating demonstration sites that serve as models for urban teaching and teacher education.

Competitive grants to schools of education and districts for developing these kinds of learning opportunities should also require evidence of teacher learning and advances in practice so that knowledge builds about how to support teachers in acquiring these much more complex teaching skills. To focus more productively on teacher performance and effectiveness, rather than merely seat time, both preparation and mentoring can be strengthened if they are guided by a high-quality, nationally-available **teacher performance assessment**, which measures actual teaching skill in the content areas. Current examinations used for licensing and for federal accountability typically measure basic skills and subject matter knowledge in paper-and-pencil tests that demonstrate little about teachers’ abilities actually to teach effectively. Several states, including Connecticut and California, have incorporated such performance assessments in the licensing process. These assessments – which can also be used as data for the accreditation process – have been found to be strong levers for improving preparation and mentoring, as well as determining teachers’ effectiveness in promoting student achievement gains. Federal support for the development of a nationally available, performance assessment for licensing will not only provide a useful tool for accountability and improvement, but it would also facilitate teacher mobility across states, which will help solve teacher shortages.

High Quality Mentoring for Beginning Teachers- Retention is at least as important to solving teacher supply as recruitment. With 30% of new teachers leaving within 5 years (and more in urban areas), the revolving door cannot be slowed until the needs for beginning teacher support are addressed. Other high-achieving countries invest heavily in structured induction for beginning teachers: they fund schools to provide released time for expert mentors and they fund other learning opportunities for beginners, such as seminars, visits to other teachers’ classrooms, and joint planning time. Such

strategies have been also been found effective in reducing beginning teacher attrition in the U.S., with rates of leaving reduced from more than 30% of beginning teachers to as low as 5% in some districts that have introduced high-quality induction programs. A critical component is strong mentoring, which includes on-the-job observations and coaching in the classroom as well as support for teacher planning by expert veterans.ⁱⁱ

Although requirements for beginning teacher induction have proliferated, with more than 30 states now requiring some kind of induction program, many are not funded and do not provide the kind of mentoring and coaching that are needed.ⁱⁱⁱ Two recent analyses of a large-scale national teacher survey revealed that, in addition to salaries and working conditions, the most important predictors of teacher's ongoing commitment to the profession are extent of preparation they have received and the quality of the mentoring and support they receive.^{iv} Federal incentives could leverage state efforts to create strong mentoring in every community. This bill provides the conditions for mentoring for beginning teachers that can reduce attrition and increase competence. If even half of the early career teachers who leave teaching were to be retained, the nation would save at least \$600 million a year in replacement costs.

Career Ladders for Teaching. The additional benefit of these and other mentoring programs is that they can be part of a **career ladder** for teachers, providing a new lease on life for many veteran teachers as well. Expert veterans need ongoing challenges to remain stimulated and excited about staying in the profession. Many say that mentoring and coaching other teachers creates an incentive for them to remain in teaching as they gain from both learning from and sharing with other colleagues.

The bill's incentives for developing career ladders in willing districts may create models that can help transform the way we organize the teaching career and keep great teachers in the profession while better using their skills. Existing compensation systems in teaching create a career pathway that places classroom teaching at the bottom, provides teachers with little influence in making key education decisions, and requires teachers to leave the classroom if they want greater responsibility or substantially higher pay. The message is clear: those who work with children have the lowest status; those who do not, the highest.

We need a different career continuum, one that places teaching at the top and creates a career progression that supports teachers as they become increasingly expert. Like the path from assistant professor to associate and full professor on campuses—or junior associate to partner in law firms—new pathways should recognize skill and accomplishment, anticipate that professionals will take on roles that allow them to share their knowledge, and promote increased skill development and expertise.

Although tying teacher advancement to performance is a desirable goal, efforts to institute versions of merit pay in education have faltered many times before – in the 1920s, the 1950s, and most recently in the 1980s, when 47 states introduced versions of merit pay or career ladders, all of which had failed by the early 1990s. The reasons for failure have included faulty evaluation systems, concerns about bias and discrimination, strategies that rewarded individual teachers while undermining collaborative efforts, dysfunctional incentives that caused unintended negative side-effects, and lack of public will to continue increased compensation.

The bill allows districts to move past these former problems by working with local teachers to develop new models that include multiple measures of performance which are carefully developed and tested. Without abandoning many of the important objectives of the current salary schedule -- equitable treatment, incentives for further education, and objective means for determining pay -- compensation systems could provide salary incentives for demonstrated knowledge, skill, and expertise that move the mission of the school forward and reward excellent teachers for continuing to teach. Rewarding teachers for deep knowledge of subjects, additional knowledge in meeting special kinds of student and school needs, and high levels of performance measured against professional teaching standards could encourage teachers to continue to learn needed skills and could enhance the expertise available within schools.^v

These initiatives generally have several features in common. All require teacher participation and buy-in to be implemented. Typically, evaluations occur at several junctures as teachers move from their *initial license*, through a period as a *novice or resident teacher* under the supervision of a mentor, to designation as *professional teacher* after successfully passing an assessment of teaching skills. *Tenure* is a major step tied to a serious decision made after rigorous evaluation of performance in the first several years

of teaching, incorporating administrator and peer review by expert colleagues. *Lead teacher status* – which triggers additional compensation and access to differentiated roles -- may be determined by advanced certification from the National Board for Professional Teaching Standards and/or other evidence of performance through standards-based evaluation systems which may incorporate evidence of contributions to student learning.

As we work to develop these new approaches to measuring teacher effectiveness for purposes of recognizing and rewarding teachers, it will be important to consider both the availability and accuracy of particular measures and the potential incentive effects of their use. For any high stakes purpose associated with personnel decision making or compensation, *multiple measures* should be used, as they are in all the systems noted earlier, since all measures give a partial picture of teacher performance and are subject to error. In addition, the system should be designed to operate so that teachers are not penalized for teaching the students who have the greatest educational needs. *Incentives* should operate to recognize and reward teachers who work with challenging students. This requires sensitivity to student and classroom characteristics in the evaluation system and ways to examine gains in learning appropriately.

Conclusion

While there are many complex elements of NCLB that will require continual attention and refinement, two important elements of the new re-authorization should be especially encouraged if we are to develop a world-class system of education. Multiple measures approaches to assessing learning – which include performance assessments of what students know and *can do* -- and multiple indicators of school performance, including graduation rates are critically important to keep the U.S. focused on developing 21st century skills for all students.

And serious investments in the teaching force – ultimately at a scale even more intensive than this bill envisions – will be the basis on which those ambitious standards can be taught and achieved. This re-authorization bill is an important start on these important agendas.

Appendix A
Assessments Around the World

Question from California High School Biology Test

5 The cell membrane of the red blood cell will allow water, oxygen, carbon dioxide, and glucose to pass through. Because other substances are blocked from entering, this membrane is called

- A perforated.
- B semi-permeable.
- C non-conductive.
- D permeable.

Questions from the 8th / 12th Grade NAEP

1. What two gases make up most of the Earth's atmosphere?
 - A) Hydrogen and oxygen
 - B) Hydrogen and nitrogen
 - C) Oxygen and carbon dioxide
 - D) Oxygen and nitrogen

2. Is a hamburger an example of stored energy? Explain why or why not.

Connecticut 9th / 10th Grade Science Assessment

Acid Rain

Student Materials

Acid rain is a major environmental issue throughout Connecticut and much of the United States. Acid rain occurs when pollutants, such as sulfur dioxide from coal burning power plants and nitrogen oxides from car exhaust, combine with the moisture in the atmosphere to create sulfuric and nitric acids. Precipitation with a pH of 5.5 or lower is considered acid rain. Acid rain not only affects wildlife in rivers and lakes but also does tremendous damage to buildings and monuments made of stone. Millions of dollars are spent annually on cleaning and renovating these structures because of acid rain.

Your Task

Your town council is commissioning a new statue to be displayed downtown. You and your lab partner will conduct an experiment to investigate the effect of acid rain on various building materials in order to make a recommendation to the town council as to the best material to use for the statue. In your experiment, vinegar will simulate acid rain. You have been provided with the following materials and equipment. It may not be necessary to use all of the equipment that has been provided.

Suggested materials:	<u>Proposed building materials:</u>
containers with lids	limestone chips
graduated cylinder	marble chips
vinegar (simulates acid rain)	red sandstone chips
pH paper/meter	pea stone
safety goggles	access to a balance

Designing and Conducting Your Experiment

- 1. In your words, state the problem you are going to investigate. Write a hypothesis using an “If ... then ... because ...” statement that describes what you expect to find and why.** Include a clear identification of the independent and dependent variables that will be studied.
- 2. Design an experiment to solve the problem.** Your experimental design should match the statement of the problem and should be clearly described so that someone else could easily replicate your experiment. Include a control if appropriate and state which variables need to be held constant.
- 3. Review your design with your teacher before you begin your experiment.**

4. Conduct your experiment. While conducting your experiment, take notes and organize your data into tables.

Communicating Your Findings

Working on your own, summarize your investigation in a laboratory report that includes the following:

- **A statement of the problem you investigated. A hypothesis (“If ... then ... because ...” statement) that described what you expected to find and why.** Include a clear identification of the independent and dependent variables.
- **A description of the experiment you carried out.** Your description should be clear and complete enough so that someone could easily replicate your experiment.
- **Data from your experiment.** Your data should be organized into tables, charts and/or graphs as appropriate.
- **Your conclusions from the experiment.** Your conclusions should be fully supported by your data and address your hypothesis.
- **Discuss the reliability of your data and any factors that contribute to a lack of validity of your conclusions.** Also, include ways that your experiment could be improved if you were to do it again.

+++++

Appendix B
Continuous Progress Toward Proficiency and Beyond:
An Index Supporting a Comprehensive Assessment System for NCLB

This paper outlines a continuous progress model of accountability for No Child Left Behind that incorporates multiple measures of school performance and operates as a growth model. The model uses an index system that assigns weights to school progress indicators. These indicators may include student achievement on a range of assessments in academic subjects, school continuation and graduation rates, and individual student growth. States establish annual targets for continuous progress both toward and beyond a high and rigorous proficiency benchmark. The proficiency benchmark and annual progress toward that benchmark can be calculated for the school as a whole and for specific subgroups within the school. The goal in this system is that 100% of schools will reach this proficiency benchmark by achieving targeted improvement increments, or Adequate Yearly Growth (AYG), each year.

The index is designed to continue the progress toward closing the achievement gap begun with the first authorization of NCLB and to support student learning and school improvement more effectively. A continuous progress index has the advantage of measuring growth and progress for students and schools at every point along a scale that includes multiple measures of student and school success. A number of states (including CA, KY, and RI) have developed similar indexes at various points in the time and have begun to develop the technology for this approach. A multiple measures approach⁶ can help schools and districts improve student outcomes more effectively because:

1. The use of multiple measures ensures that attention will be given to a comprehensive academic program and a more complete array of important learning outcomes;
2. A multiple measures approach can incorporate assessments that evaluate the full range of standards, including those addressing higher-order thinking and performance skills;
3. Multiple measures provide accountability checks and balances so that emphasizing one measure does not come at the expense of others (e.g. boosting test scores by excluding students from school);
4. A multiple measures index can provide schools and districts with incentives to attend to the progress of students at every point on the achievement spectrum, including those who initially score far below or above the test score cut point labeled “proficient.” It can encourage schools to focus on the needs of low achievers, students with disabilities, and ELL students, using assessments that measure gains from wherever students begin and keeping them in school through graduation.

One of the central concepts of NCLB’s approach is that schools and systems will organize their efforts around the measures for which they are held accountable. Because

attending to any one measure can be both partial and problematic, the concept of multiple measures is routinely used by policymakers to make critical decisions about such matters as employment and economic forecasting (e.g., the Dow Jones Index or the GNP), as well as admissions to college. Successful businesses use a “dashboard” set of indicators to evaluate their health and progress, aware that no single indicator is sufficient to understand or guide their operations. This approach is designed to focus attention on those aspects of the business that describe elements of the business’s current health and future prospects, and to provide information that employees can act on in areas that make a difference for improvement. So, for example, a balanced scorecard is likely to include among its financial indicators not only a statement of profits, but also cash flow, dividends, costs and accounts receivable, assets, inventory, and so on. Business leaders understand that efforts to maximize profits alone could lead to behaviors that undermine the long-term health of the enterprise.

Similarly, a single measure approach in education can create unintended negative consequences or fail to focus schools on doing those things that can improve their long-term health and the education of their students. Although No Child Left Behind calls for multiple measures of student performance, the implementation of the law has not promoted the use of such measures for evaluating school progress. The focus on single, often narrow, test scores in many states has begun to narrow the curriculum and reduce access to education for the most vulnerable students. A multiple measures approach that incorporates well-balanced set of indicators would support a shift toward “holding states and localities accountable for making the systemic changes that improve student achievement” as has been urged by the Forum on Education and Accountability, a consortium of 120 civil rights and education organizations. Moreover, a multiple measures approach incorporating a set of research-based indicators would broaden the scope of accountability.

What Measures Could be Used in the Proficiency Index?

The system would use a set of measures, combined in a weighted index that establishes an annual score for each school and sets an annual proficiency target for improvement as the basis for determining Adequate Yearly Progress (AYP). The index would enable a baseline to be set for subgroup accountability that would encourage significant growth for each group, based on its unique starting point. Measures should represent important school outcomes and factors that are associated with school improvement and student opportunities to learn. These could include:

- Scores on student assessments, including but not limited to assessments in reading and mathematics, and including measures of student year-to-year gains;
- Measures of student participation and progress in school, including data about attendance, student grade-to-grade progression and continuation through school, suspension rates, and graduation rates;
- Additional quantitative measures of school learning conditions and effectiveness, including, for example, data about student participation in

rigorous coursework (e.g. Advanced Placement or college courses at the high school level), and school supports.

The rationale for including indicators of learning conditions is to encourage schools to attend to aspects of school quality that influence students' opportunities to learn. In the business world, these kinds of measures are called leading indicators, which represent those things that employees can control and improve upon. These typically include evidence of customer satisfaction, such as survey data, complaints and repeat orders; as well as of employee satisfaction and productivity, such as employee turnover, project delays, evidence of quality and efficiency in getting work done; reports of work conditions and supports, and evidence of product quality. Educational versions of these kinds of indicators are available in many state accountability systems. For example, Rhode Island measures school learning conditions in part through an annual survey to all students, teachers, and parents that provides data on "Learning Support Indicators" measuring school climate, instructional practices, and parental involvement. In addition, Rhode Island, like some other states, conducts school quality reviews, not unlike the Inspectorate system that is used in the UK and many other countries. These kinds of reviews can examine teaching practices and the quality of learning opportunities.

This "School Growth" model may incorporate an "Individual Growth" model for tracking the progress on assessments of individual children. Incentives for individual student growth, such as those federally authorized for nine states under the Growth Model Pilot Program, are more technically intricate than school progress models in that they require 1) tracking systems for individual students over time and 2) "vertically linked" assessments that report meaningful progress across grades. The proposed continuous progress index does not require individual growth models in the near term. However, states with existing infrastructure for individual growth measurement are strongly encouraged to incorporate individual growth into their indices.

In addition, one could consider requiring a separate *Opportunity-to-Learn Index*, reported for each school, which would provide data about school capacity, such as evidence about teacher and other staff qualifications, availability of learning materials (books, computers, etc.), curriculum access, and funding. This would be a powerful lever for encouraging states and districts to attend to equity concerns, as it would make much more visible the kinds of resources needed and available to meet standards in each school.

How Would the System Work?

The index would be developed by states and approved by the Department of Education, using a fully transparent peer review process, according to specific criteria (listed below). The approval process should ask states to provide a rationale for their set of indicators and weights and should evaluate how the state proposal will support the proposed school improvement process for schools failing to make progress. Ongoing review of state plans should evaluate how well the state's approach is working to support improvement in

schools over time, providing a basis for holding states accountable. Each state's system would include the following components:

- 1) A weighted set of indicators – A set of assessment scores, student progress indicators, and other measures to which weights are assigned reflecting the relative importance of each component.
- 2) An index to measure current status and growth at the district, school, and subgroup levels - A method of combining indicators to produce a score on an index that can be tracked each year for the school and for subgroups within the school, and aggregated to the district level.
- 3) A proficiency benchmark for all schools – A defined numeric score on the index which represents a high level of achievement on assessments of student learning, combined with a high rate of student participation in school (e.g. school progression and graduation rates) and other indicators of school quality. This serves as the goal toward which all schools are working as assessed by measures of annual progress. The proficiency benchmark should be set at a level that represents a high, achievable target, such as that currently attained by the top quartile of schools in the state.
- 4) A means to set annual growth targets toward the proficiency benchmark – A formula, informed by research, that sets a target score on the index for each school each year based on the prior year's score and the expectation for adequate yearly progress.

Criteria

The federal government would specify criteria a state system would need to meet. These are policy decisions that would be developed to reflect federal priorities. For example:

1. The index must include multiple measures of student learning that are aligned to the state standards, including but not limited to assessments of student learning in reading and mathematics. (These could include assessments in subjects beyond reading and mathematics, and a mix of measures, including performance assessments, within or across specific subject areas.) Student learning would also be evaluated by student progression through graduation from school, including a standard graduation rate calculation. Graduation rates will be calculated as described below. Both 4-year and 5-year graduation rates will be included in the index.
2. Reading and mathematics assessments should comprise at least 50% of the total value of the index and, for secondary schools, progression and graduation rates should count at least 20%. [Note: these illustrative weights represent one view of educational value, and could be adjusted to reflect different priorities. Some would argue, for example, that graduation rates should be equivalent to

- achievement data. Others might argue that assessments should count for more than 50% of the total.]
3. The index may also include additional evidence about learning conditions, including student access to rigorous coursework, which may comprise up to 20% of the index.
 4. The index must be technically sound, relying on measures that are reliable and valid for the students who are being evaluated and for the purposes for which they are used. Test measures should give students a fair chance to show what they know and have learned. [Testing standards are outlined below.]
 5. The index should assess learning and reflect changes across the full distribution of achievement, evaluating and valuing growth among low-achieving, average-achieving, and high-achieving students. Wherever possible, the index should measure growth using student-level longitudinal performance data, as well as absolute scores. Student gains should be weighted more heavily than absolute scores.
 6. The index must provide disaggregated index scores for subgroups defined by race / ethnicity, socioeconomic disadvantage, disability status, and language proficiency.
 7. The index should reflect the achievement and growth of all students, including English language learners and those with special needs. Where alternative measures are needed to validly assess the achievement of ELL and SWD students, these should permit the assessment of gains in learning, and such gains should be factored into the index.
 8. For purposes of reporting gains in subgroup scores, students who were initially recorded as part of a subgroup shall have their index scores reported as part of the subgroup throughout their time in the school so that progress for the subgroup can be appropriately tracked. This designation for reporting purposes shall not determine student program placements.
 9. Growth targets on the index must be set to assure substantial annual progress toward the proficiency benchmark. [For example, the annual growth target for a school -- and for each significant subgroup within the school -- might be 5 percent of the difference between the school's current index score and the proficiency benchmark.] Determination of adequate progress for purposes of intervention will include three years of data.

How Might the Index Operate?

The example below shows different indexes for two hypothetical states. The first is based on the law's minimal requirements; the second includes a fuller range of indicators. Both

states have established a Proficiency Index which measures schools' performance. This index is a numeric scale that ranges from a low of 200 to a high of 1000. Measures are weighted as follows:

Index Weights, Elementary / Middle Schools, Grades 3-8

Content Area	Index Weights State A	Index Weights State B
Assessments (Total)	0.50	0.60*
English-Language Arts	0.25	0.15
Mathematics	0.25	0.15
Science		0.15
History-Social Science		0.15
Student Participation and Progression	0.50	0.30
Student attendance	0.20	0.10
Grade-to-grade progression / continuation	0.30	0.15
Suspension rates		0.05
Additional Measures of School Learning Conditions / Outcomes	NA	0.10

Index Weights, High School, Grade Levels 9-11

Content Area	Index Weights State A	Index Weights State B
Assessments	0.50	0.70*
English-Language Arts assessments	0.20	0.20
Mathematics assessments	0.20	0.20
Science assessments	0.10	0.15
History-Social Science		0.15
School Participation, Progression, and Graduation	0.50	0.25
Progression and graduation rates	0.45	0.15
Attendance	0.05	0.05
Suspension rates		0.05
Additional Measures of School Learning Conditions / Outcomes		0.05

* In some content areas, the set of assessments includes state-developed, curriculum-embedded performance assessments that are factored into the score. For example, in English language arts the on-

demand state test is weighted 0.10 and the set of performance assessments, which include more extended writing tasks scored using a state rubric, is weighted 0.05.

A school's score on the Index is an indicator of a school's performance level. The statewide Proficiency Benchmark for all schools is 750, a number based on calculations for proficient performance on the combined assessments, a targeted 90% graduation rate, and appropriate benchmarks for the other indicators (e.g. 95% attendance, 90% grade-to-grade progression rates, 5% suspension rates, etc.). A school's growth is measured by how well it is moving toward or past that goal. A school's score in the prior year is subtracted from its score in the current year to determine how much the school improved in a year. The school target for the next year is 5% of the difference between its current score and the Proficiency Benchmark of 750.

Calculation

The index is calculated as the weighted average of student scores across content areas + the weighted average of each other indicator (e.g. graduation rate, attendance rate). Test scores are scaled so that 750 is the equivalent of a “proficient” performance level.

Testing Standards

Testing standards will draw on the Joint Psychological Standards for Testing developed by the APA, AERA, and NCME to outline standards for test use, addressing:

- Technical quality of assessment
- Comprehensive alignment to state standards
- Use of criterion-reference assessments that allow measurement of student progress on the standards and school progress in a benchmarked system
- Freedom from bias
- Validity and reliability issues, including validity for the specific students and purposes for which assessments are to be used
- Construction and use of appropriate measures for specific populations, including how to use appropriate assessments that allow measurement of gain scores at all points of the achievement spectrum for students with disabilities, aside from the small percentage exempted from the accountability system, and how to use and chart gains for ELL students on the assessments that are most valid for them as they acquire English language proficiency.
- Treatment of scores in the accountability system

Endnotes

ⁱ For recent studies examining the increases in dropout rates associated with high-stakes testing systems, see Advocates for Children (2002). Pushing out at-risk students: An analysis of high school discharge figures -- a joint report by AFC and the Public Advocate. <http://www.advocatesforchildren.org/pubs/pushout-11-20-02.html>; W. Haney (2002). Lake Wobegone guaranteed: Misuse of test scores in Massachusetts, Part 1. Education Policy Analysis Archives, 10(24). <http://epaa.asu.edu/epaa/v10n24/>; J. Heubert & R. Hauser (eds.) (1999). High stakes: Testing for tracking, promotion, and graduation. A report of the National Research Council. Washington, D.C.: National Academy Press; B.A. Jacob (2001). Getting tough? The impact of high school graduation exams. Education and Evaluation and Policy Analysis 23 (2): 99-122; D. Lilliard, & P. DeCicca (2001). Higher standards, more dropouts? Evidence within and across time. Economics of Education Review, 20(5): 459-73; G. Orfield, D. Losen, J. Wald, & C.B. Swanson (2004). Losing our future: How minority youth are being left behind by the graduation rate crisis. Retrieved September 10, 2007 from: <http://www.urban.org/url.cfm?ID=410936>; M. Roderick, A.S. Bryk, B.A. Jacob, J.Q. Easton, & E. Allensworth (1999). Ending social promotion: Results from the first two years. Chicago: Consortium on Chicago School Research; R. Rumberger & K. Larson (1998). Student mobility and the increased risk of high school dropout. American Journal of Education, 107: 1-35; E. Rustique-Forrester (in press). Accountability and the pressures to exclude: A cautionary tale from England. Education Policy Analysis Archives; A. Wheelock (2003). School awards programs and accountability in Massachusetts.

ⁱⁱ A number of studies have found that well designed mentoring programs improve retention rates for new teachers along with their attitudes, feelings of efficacy, and their range of instructional strategies (California Commission on Teacher Credentialing, 1992; Karge and Freiberg, 1992; Kolbert and Wolff, 1992; Darling-Hammond & Sykes, op. cit.; Luczak, op. cit.)

ⁱⁱⁱ National Commission on Teaching and America's Future, *No Dream Denied*. Washington, DC: Author, 2003.

^{iv} Ingersoll, 1997b; Luczak, 2005.

^v There are a few initiatives that have been developed which share some common features upon which these new career ladders may be built. These include the **Procomp system in Denver, CO**, a locally-designed system of teacher compensation based on knowledge, skills, and performance, which incorporates compensation for additional knowledge and skills, evidence of performance, and new roles and responsibilities. (For more detail, see <http://denverprocomp.org>.) and innovative career ladders in **Cincinnati, OH** and **Rochester, NY**, which have been in place since the 1990s. These incorporate mentoring for beginners by lead teachers who have reached the top level of the career ladder, as well as other responsibilities for lead teachers. More recent efforts have been initiated in **Helena, Montana** and **Portland, Maine** and, in the only current state system, **Minnesota** (for more information, see http://www.educationminnesota.org/index.cfm?PAGE_ID-15003). In addition the **Teacher Advancement Project** provides a career ladder that rewards evidence of knowledge and skill, as well as performance, and incorporates extensive professional development that occurs during shared teacher time in schools that are redesigned to provide regular learning opportunities.

⁶ "Multiple measures" refers to sources of evidence about student learning that a) use different methods for demonstrating achievement and b) provide multiple opportunities to demonstrate achievement that are accessible to students at varying levels of proficiency. Such measures may include locally-managed assessments that are reliable and valid as well as standardized state assessments.