

Committee on Education and Labor
Subcommittee on Early Childhood, Elementary and Secondary Education Hearing
“How NCLB Affects Students with Disabilities”
March 29, 2007, 10:30 a.m, Room 2175 Rayburn House Office Building

Thank you Chairman Kildee, Ranking Member Castle, and all the Members of the Subcommittee for inviting me to testify this morning.

I am the Principal Investigator of an Office of Elementary and Secondary Education (OESE) funded Enhanced Assessment Grant (EAG) contract with the state of Georgia, and Co-Principal Investigator of another OESE EAG contract with the state of New Hampshire, both of which are housed at the National Center on Educational Outcomes at the University of Minnesota. These projects are focused on so-called “gap assessment issues” and on alternate assessments. I also am the Co-Principal Investigator and technical assistance team leader for a National Technical Assistance Center on Assessment for Children with Disabilities, funded by the Office of Special Education Programs (OSEP), also housed at the National Center on Educational Outcomes (NCEO). In that role I provide technical assistance to all the states and territories on inclusive assessment and accountability systems. Finally, I participate in research on technical issues of alternate assessment through a subcontract with the National Alternate Assessment Center (NAAC), a research center funded by OSEP, and housed at the University of Kentucky. However, in my testimony this morning, I am representing myself, and not the University of Minnesota or the multiple projects on which I work.

The Federal role in public education is a frequent topic of news articles and blogs these days, but the idea of a strong public commitment to ensure that all the children of all of the people are educated to be effective members of our democracy is not new. Our noble experiment in government by the people and for the people has everything to do with our discussion today. The civil rights era during the mid 20th century opened the doors to the same public education system for all students. It was at that time that the system was opened for many of the children who still are on the wrong side of the achievement gap. Many of these students, including most students with disabilities, had access to the school room by the 1970s, but did not have access to the same rich curriculum and effective instruction as their peers. Our discussion today is on one level simply a continuation of three decades of work. That we are still seeing persistent gaps is the bad news – but the good news is that in the last 6 years, since the passage of the No Child Left Behind Act, we have learned a lot about why all children have not been achieving at high levels, and we have seen examples around the country of what can happen when we change our practices.

I will focus my remarks on the standards, assessment, and accountability components of the No Child Left Behind Act, and how they have affected students with disabilities. I hope to share examples in three areas:

First, I will share data from schools where they have changed practice, making use of assessment data and the pressure of adequate yearly progress as an opportunity for reform

so that all students can achieve the challenging standards set for all students. These examples demonstrate that the long-term goal of full proficiency is not an illusion, and at a minimum, we can not at this point accept the argument that we should accept far less.

Second, I will share what we have learned about assessing the academic performance of ALL students during the past decade, including those students who may not take a “typical” or “on average” path through the content. Think of students who are deaf or students who have barriers to specific processing skills – we have learned, and are learning, so much about HOW students learn the full range of challenging content because of the push from NCLB to understand how well they HAVE learned. This push has resulted in improved understanding and design of standards-based assessments.

And that, in turn, leads to better understanding and design of accountability systems based on these assessments, which is my third point. It is in this area that we have the most to learn still, but improving the quality of the assessments will take us part of the way to that outcome.

After waiting for three decades for substantive progress in educating ALL children, I am here to say that NCLB has resulted in public clarification of and support for high expectations for all students, improvement of assessment design and use, and the spotlighting of effective and ineffective practices in public schools, for all students. That makes it the most powerful lever we have found in the past three decades to move students with disabilities, along with other students who have been affected by the achievement gap, from mere ACCESS to public schools to ACHIEVEMENT. I will elaborate by focusing on the three topics.

CHALLENGING STANDARDS FOR ALL STUDENTS: First, let me address what is sometimes called the “existence” proof of student achievement. Anecdotal stories have been building for several years, and many states have instituted formal procedures to use assessment and accountability data to identify schools where reforms are yielding very high achievement for students with disabilities. We have also had a few formal studies of what is occurring in schools where test scores are higher for students with disabilities. The informal and formal studies have consistently found that schools where students with disabilities are achieving at high levels share these characteristics, as summarized in one study: (1) a pervasive emphasis on the curriculum and alignment with the standards, (2) effective systems to support curriculum alignment, (3) emphasis on inclusion and access to the curriculum, (4) culture and practices that support high standards and student achievement, (5) well-disciplined academic and social environment, (6) use of student data to inform decision making, (7) unified practice supported by targeted professional development, (8) access to resources to support key initiatives, (9) effective staff recruitment, retention, and deployment, (10) flexible leaders and staff that work effectively in a dynamic environment, and (11) effective leadership (Donahue Institute, 2004).

The February 2007 edition of *Educational Leadership* (the professional journal of the Association for Supervision and Curriculum Development - ASCD) features a

longitudinal study in Rhode Island. The study, completed through the local affiliate of ASCD, looked at “closing of the gap” and found that 100 of Rhode Island’s 320 public schools had shown dramatic closing of the gap for the students with disabilities subgroup over the 2001-2004 testing period. They followed up with a survey of these schools, and include summaries of the specific reform strategies that these schools incorporated. The author notes that:

The diverse populations of schools that have successfully reduced the achievement gap suggest that strategies are applicable and effective in a variety of settings. For example, Richmond Elementary is a rural school that enrolls 500 students, with a minority population of approximately 2 percent. Duterple, on the other hand, has a minority population of more than 28 percent and is defined as “urban ring.” (Hawkins, 2007, p. 63).

The leaders of these schools ensured that all teachers had the support, skills, tools, and strategies they needed to effectively teach the same challenging content to all children. These and the other common practices identified in these 100 successful schools mirror the findings of the Donahue Institute, and the author concludes:

The Rhode Island ASCD affiliate study proved that students with special needs can achieve high standards when schools address learning needs. Successful schools had strong leadership and incorporated effective practices that promoted a responsive learning environment. Most important, they were committed to ensuring the success of each student. (Hawkins, 2007, p. 63).

If we have existence proofs of dramatic acceleration of student achievement following best practice interventions, it raises the question of why the letters to the editor pages are so full of teachers, and even parents, decrying the expectation that all children learn to high levels. Why have educational professionals so resisted actually teaching students with disabilities the challenging content, and expecting them to learn it? Part of the answer to this rests in centuries of fear and bias, or pity and caretaking toward people with disabilities, or for that matter, any people who are different from the typical.

HIGH QUALITY STANDARDS-BASED ASSESSMENTS: Now I will turn to the second point – once all students have been taught well the challenging curriculum, how do we ensure that we can show what they know appropriately on assessments? If our academic assessments are not sensitive to capture what students actually have learned, our system will harm rather than help our efforts to ensure all children succeed. This second topic does not have the same public appeal as telling the student and school success stories, but is essential for understanding how NCLB has and can continue to increase our understanding of how all children can achieve proficiency. Let me start with what the data from fully inclusive assessments have told us thus far, but then let me step back to a discussion of what a “good” assessment is, based on work from the National Research Council’s Committee on Assessment.

NCEO surveys of States over the past decade have recorded state staff perceptions of changes occurring in their districts and schools. They speak of the improvements in the performance of their students, attributing the improvements to clear assessment

participation policies, alignment of IEPs with standards, improved professional development, development and provision of accommodation guidelines and training, increased access to standards-based instruction, and improved data collection (Thompson, Johnstone, Thurlow, & Altman, 2005). Analyses of publicly reported assessment data since 2000-2001 show improvements in the transparency of data for students with disabilities, both for participation and for performance (Thurlow, Quenemoen, Altman, & Cuthbert, in press). For example, NCEO's identification of states with clear participation reporting to the public for students with disabilities showed only 5 states in 2000-2001, but 20 states in 2004-2005. These data also showed large increases in participation percentages across time for most states. Data on performance showed similar changes – more states with clear transparent reporting, and increases in performance across years (see example from Massachusetts in Appendix A). Data from the National Assessment of Educational Progress (NAEP) support these data – overall there has been an increase in the performance of students with disabilities (see Appendix B).

Assessment literacy basics – norms vs. standards: So the news is generally good. We have had dramatic success in some schools, but dramatic resistance to change in schools often just down the highway in the same state or even district. Some of the resistance comes from misunderstanding of what standards-based testing is meant to measure. Most educators, and for that matter, members of Congress, have taken large-scale achievement tests throughout their schooling years. These tests typically were built on the measurement models of the 20th century – norm referenced tests, or NRTs, designed to sort us into bell-shaped curves on some kind of ability distribution. These tests have not given us precise information at the individual level that will tell us who has deep and enduring understanding of important knowledge and skills in math or in reading. They have been well suited to give us a general sense of how a group of students is performing, as well as serving to sort out huge populations for purposes like army personnel assignments or admission to elite colleges. We have a century of development on measurement models that work well for these purposes.

These norm-referenced tests are designed for a very different purpose and use than tests that would tell us what groups of students know compared to a well-defined standard, or a criterion (or standards-based) referenced test (CRT). Not all states are using tests that do this well, but the testing industry is making progress toward better test design for the purpose of comparison to well-defined content and achievement standards and not to normal curves. Still, over a century of norm-referencing testing designed to distribute students along a normal curve has affected the perceptions of teachers, parents, and the public – and has resulted in a popular belief that on any skill taught, we can expect and should see on tests that half of the students are “below average.”

Garrison Keillor has made use of these misconceptions in his signoff from Lake Wobegon, not far from my home, “where all the children are above average.” If the middle class, Lutheran students of Lake Wobegon are taking a norm-referenced test, that is very probably true, for a variety of complex reasons. If they are taking a high quality criterion referenced test based on challenging content and achievement standards, then

there is not an “average” to describe, only relative distance from the standard. Then, we would hope, students who have been taught well will demonstrate proficiency on the standards. This fundamental understanding of testing is missed by much of the public (and many educators) in discussions of No Child Left Behind. (For more information on distinctions between NRT and CRT testing purposes and uses, see CCSSO Handbook for Professional Development in Assessment, Sheinker & Redfield, 2000).

So if there is a widely accepted (but erroneous) assumption that there will ALWAYS be students who do poorly on tests (i.e., below average), then it is pretty tempting to predict which students will end up on the bottom. Have you ever heard educators or members of the public say, “Well of COURSE they don’t do well on the tests, they have disabilities!” Once those expectations are entrenched, they play out in insidious ways. The literature on the effects of teacher expectations on student achievement is deep and strong – and alarming given so many educators seem to believe that students with disabilities cannot learn well.

On the other hand, if you understand that you teach all the children the same content, tailoring services, supports, and specialized instruction for some children to be sure they master the content, and then you test to ensure that they have indeed learned it, then your views of testing are very, very different. And it leads to a very important point, and one that has been advanced by NCLB requirements. In order to measure standards-based learning for all students, including those who do not “fit” the measurement models we have developed the past century, we have to rethink some of the assumptions we have made about testing.

How do we know what ALL students know? Fortunately, we are addressing these issues after the National Research Council’s Committee on Assessments worked on the problems of educational assessment, resulting in the book, *Knowing What Students Know: The Science and Design of Educational Assessment* (NRC-Pellegrino, Chudowsky, & Glaser, 2001). We have made much progress. The Winter 2006 (just published) journal of the National Council on Measurement in Education (NCME), *Education Measurement: Issues and Practice*, devoted an entire issue to what we are learning. The editor of the special topic journal states in the opening: “These papers reflect the current state of the art; even the *Standards for Educational and Psychological Testing* do not address assessment of these subgroups [LEP and some SWD] fully...A psychology of educational achievement testing would have to address special considerations for these student subgroups.” (Ferrara, 2006, p. 3).

In our work with the New Hampshire Enhanced Assessment Grant the past two years, we have been able to partner with multiple state leaders, and with measurement, curriculum, and special education experts to understand how the current state of the art reflected in the Committee’s work applies to students with disabilities. A few basic principles from the NRC work help illustrate the key findings, and why they are important to ensuring we truly know what students know.

The NRC work builds from a simple triangle that represents the process of reasoning from evidence.

The corners of the triangle represent three key elements underlying any assessment: [1] A model of student cognition and learning in the domain, [2] a set of beliefs about the kinds of observations that will provide evidence of students' competencies, and [3] an interpretation process for making sense of the evidence" (Pellegrino et al., 2001, p. 44).

In applying this to students with disabilities, we have asked similar questions: Who are the students – what are their characteristics – and how do they build competence in the domain? What observations of their learning will illustrate how well they have learned? What inferences can we make about their learning?

Who the students are, and their learning characteristics, varies greatly among students with disabilities. See Appendix C for a summary chart of categorical distribution of students with disabilities. Based on a recent summary by NCEO's Director, Martha Thurlow:

Most students with disabilities (75% altogether) have learning disabilities, speech/language impairments, and emotional/behavioral disabilities. These students, along with those who have physical, visual, hearing, and other health impairments (another 4-5%), are all students *without intellectual* impairments. When given appropriate accommodations, services, supports, and specialized instruction, these students (totaling about 80% of students with disabilities) can learn the grade-level content in the general education curriculum by going around the effects of their disabilities, and thus achieve proficiency on the grade-level content standards. In addition, research suggests that many of the *small* percent of students with disabilities *who have intellectual impairments* (less than 2% of the total population of all students, which is *less than 20% of all students with disabilities*), can *also* achieve proficiency when they receive high quality instruction in the grade-level content, appropriate services and supports, and appropriate accommodations.

The key to this description is to think about how these students build competence in the academic domains. If we have "existence proofs" that these students can learn the same challenging content as their peers, would you guess that they may vary in their "paths to proficiency?" There are some assessments that assume that all students take the same path to proficiency. What we realized is that there are reasons to doubt that assumption for this highly varied group of students with disabilities. We have limited good data to understand how they vary. In the early days of implementation after P.L.94-142, special education was built on a deficit model – with the goal of overcoming the identified problem in children, or caring for those for whom the deficits were severe – and was not based on a curriculum model. Thus, special education trained people generally have a very limited understanding of content domains. This is compounded by the history in special education of seeing the "general curriculum" as a tightly ordered list of tasks, from low to high, K-12, and every skill in sequence is necessary for subsequent ones.

Although all subspecialties or philosophical positions in education do not agree to what degree this task-analyzed view of curriculum is or is not true, there IS agreement among

curriculum professional organizations – as illustrated by mathematics and English teachers’ professional organizations’ own standards – that building competence in the domain is not a lock-step, task-analyzed march to proficiency in a narrow range of basic skills. Not all lower level skills or knowledge are necessary in order to master more challenging content, bigger ideas, or problem-solving– the engaging and motivating content beyond basic skills. We are just beginning to understand how concepts develop, with the NRC Committee work suggesting, “Not all children learn in the same way and follow the same paths to competence” (Pellegrino et al., p. 103), and they are referring to all children. How this applies to students with disabilities is even less well understood.

What we have learned about what we know – and do not know – is striking. Tests – and the statistical models that we have trusted in the past – are only as good as our understanding of who the students are and how they build competence in each domain – not just in the aggregate across tens of thousands, but the groups who by and large may take an atypical path to competence. And they are only as good as the curricular philosophy and quality of design that is represented by the items on the test. Are assessments up to the task? What are the underlying assumptions of the statistical models? Are those assumptions met for all subgroups? As pointed out in the NCME journal cited above, we are forced to rethink validity frameworks now that we have included ALL students.

Do current assessments work for students with disabilities? All assessments are not equal in their capacity to show well what all students know. Still, state assessment offices have in many cases pushed the testing industry and the measurement field to examine long-established practices that may, in fact, obscure what some students know and are able to do. This is especially true in states where assessment, curriculum, and special education offices have collaborated on conceptualizing the assessment system to ensure all students can show what they know. These states typically begin with a deep understanding of the full range of the academic domains to be assessed, then carefully consider who the students are who are deemed “difficult to assess,” and how these students show real growth in knowledge and skills of the domain. These deliberations, often done in the absence of a research base, depend on carefully looking at actual student work that has resulted from effective teaching of the content. Only then can a state decide how to design the assessment, and what can be learned from results on the assessment. In many states, we have found that the National Research Council assessment triangle is especially important to consider in the design, administration, and use of assessments where “atypical learning paths” are to be measured. What we realize however, is that these assessments carefully designed to capture how students learn and show what they know give us far better information on learning for all students, not just those with special needs.

Regular assessment, with or without accommodations. The vast majority of students with disabilities take the regular assessment, with or without accommodations. These tests are increasingly developed with universal design elements in mind, resulting in more accessible tests for all students. For example, students with disabilities are considered from the beginning in the development of test design, so that item features are not

included that assess disabilities rather than the construct; test item developers have been trained, and bias and sensitivity reviews consider disability issues in connection with content and construct targets. These design features result in more accurate scores that reflect actual student knowledge and skills, and not extraneous factors. Universal design does not mean that accommodations are no longer necessary.

Our knowledge base on the effects of accommodations on the content being measured is growing, but the complexities of the research are considerable for the most challenging content and student combinations. Most accommodations are widely accepted as not changing what is being measured, but there are accommodations that are controversial, or that are considered modifications. Modifications are changes in administration that do indeed change what is measured, and change the meaning of the test results. For a very few students, these modifications are the only way they can interact with portions of the test, so states are working hard to develop strategies to capture what the student can do while ensuring that standards are not lowered. States are struggling with policies that retain high expectations for all students as well as integrity in measurement and meaning of the test. A few states have made and defended sometimes controversial decisions on these issues, but they also require close monitoring and accountability for schools where these accommodations are selected for students. It is a delicate balance between access and inadvertently lowering expectations, and standards. (see NCEO teleconference materials from accommodations series, 2005-2006:

<http://education.umn.edu/nceo/Teleconferences/tele11/default.html>

<http://education.umn.edu/nceo/Teleconferences/tele12/default.html>

<http://education.umn.edu/nceo/Teleconferences/tele13/default.html>

All the attention on accommodations policies for testing has had a powerful consequence. It has drawn much needed attention to what has been limited understanding and quality of use of accommodations for instruction or assessment (Lazarus, Thompson, & Thurlow, 2006, <http://www.epri.org/PDFs/IB7.pdf>). The CCSSO sponsored ASES SCASS collaborative of about 20 states developed a training module, now included in the OSEP toolkit, to ensure states had high quality training materials for accommodations use (Thompson, Morse, Sharpe, & Hall, 2005).

http://www.osepideasthatwork.org/toolkit/accommodations_manual.asp

The training module, which is designed to be adapted by states to encompass their own policies, provides training to build teacher and IEP team capacity on these five steps:

- Expect students with disabilities to achieve grade-level academic content standards.
- Learn about accommodations for instruction and assessment.
- Select accommodations for instruction and assessment for individual students.
- Administer accommodations during instruction and assessment.
- Evaluate and improve accommodation use

The first bullet is essential. Accommodations decisions must be made individually, based on the how the student builds competence in specific knowledge and skills of the academic domain. As pointed out in the opening of this written testimony, special educators often do not have strong content knowledge – and the accommodations

decision process breaks down when there is limited understanding or misunderstanding of the content to be taught and assessed. The key questions that are difficult to make decisions about without content knowledge are: Does the academic content being learned or assessed change because of the proposed accommodation? Will the use of this accommodation prevent the student from mastering the content fully, and thus possibly limit future understanding of related content? Strengthening the link between accommodations decisions and deep understanding of the content domains remains a barrier to effective use of accommodations in instruction and in assessment. Special educators who do not have content expertise must form strong partnerships with content experts. Those with content expertise must be reminded that the purpose of accommodations is to maintain the validity of the grade-level construct that is being measured. Fortunately, many special and general educators are working in tandem to design effective instruction that has been shown to be highly effective.

The bottom line of NCLB accommodation testing requirements is that students who need accommodations are more likely today than before NCLB to be receiving needed accommodations – during testing and during instruction. Prior to NCLB, high estimates indicated that the use of testing accommodations was about 53% for elementary schools and 44% for middle and high schools (Thompson & Thurlow, 1999). Today those percentages are 65% for elementary schools, 64% for middle schools, and 62% for high schools (Thurlow, Moen, & Altman, 2006). According to the authors, “We do not know, as yet, what the percentages *should* be, but having them closer to the same across school levels is a good sign.”

There are students with disabilities who participate in the regular assessment, with or without use of accommodations, whose scores are very low. Low scores are NOT a sign that the test doesn't work – the test results tell us how well students are doing so something can be done about it. Many states have carefully analyzed who the “persistently low-performing” students are on their general assessment (e.g., see NCEO teleconference, specifically the NCIEA and Colorado studies, <http://education.umn.edu/nceo/Teleconferences/tele11/default.html> and also studies from the New England Compact Enhanced Assessment Grant, 2007). In all cases, they have found that persistently low performing students are NOT all students with disabilities, and generally include other subgroups such as minority or disadvantaged students. In addition, in all cases where they have investigated what instructional opportunities these low performing students had, they found that they were not being taught the content on the test. In the Colorado study, they found that students were not being given accommodations that would help them learn the content and then show what they know.

Alternate assessments.

There is a small group of students who cannot show what they know on the regular assessment, even with use of accommodations, and alternate assessments have been developed for many of these students. There are two options for alternate assessment allowed in NCLB law and regulation. One is the alternate assessment based on grade-level achievement standards and the other is the alternate assessment based on alternate achievement standards.

These alternate assessment options are meant to measure how well students know and can do their enrolled grade content standards, but the options DIFFER as to HOW WELL and even HOW MUCH the student knows and can do on that grade level content. They also differ in who the students are for whom the assessments are intended. With the rapid evolution of our understanding of how students build competence, especially students who need alternate assessments, there are many misconceptions and erroneous assumptions related to these alternate assessments. It is important to build a common understanding of the distinctions so the goals of reform are not derailed by these misunderstandings. These alternate assessments provide a critical role in ensuring that we can truly obtain accurate measures of the knowledge and skills of all students with disabilities. It is important to take a bit of time to summarize these approaches to assessment and what we know about them.

Alternate assessments based on grade-level achievement standards are meant to assess the SAME content with the SAME definition of “how well and how much” of the content as is measured by the regular assessment. There are very few, if any, such alternate assessments in place in states now, in part because of the very difficult measurement challenge of showing comparability to the regular assessment. This is an area we hope can develop over time.

Alternate assessments based on alternate achievement standards are meant to assess the SAME content with a DIFFERENT definition of “how well and how much” of the content as is measured by the regular assessment. These assessments are for students with significant cognitive disabilities. Many of the projects and states I work with are gradually developing better understanding of these definitions as we see more and more evidence of student work to guide us. The bottom line for these assessments is that in emerging consequential validity studies in multiple states, we are finding that students with significant cognitive disabilities have dramatically increased in their access to the general curriculum because of the NCLB requirements for these assessments. There have been reports of dramatic increases in other valued outcomes for students with significant cognitive disabilities concurrent with their participation in accountability, specifically, increased use of assistive technology, which in turn increases level of independence, increased implementation of inclusive settings, and increased interaction with typical peers (Horvath, Kampher-Bohach, Kearns, 2005; Kampfer, Horvath, Kleinert, & Kearns, 2001; Kearns, Kleinert, Kennedy, 1999; Turner, Baldwin, Kleinert, & Kearns, 2000).

The achievement of these students on the content is very different from their inclusive classroom peers, but the evidence of their work is compelling: these students are able to learn academic content with reduced complexity, breadth, and depth clearly linked to the same grade level content as their peers. Researchers and practitioners are working side by side to capture the nature of the linkages to the grade-level content, but the evidence of their learning is startling, given that we have not given them access to this content in the past.

Despite the very dramatic and positive outcomes that have been achieved in many locations for students with significant cognitive disabilities, the changes have been challenging and painful for some teachers and other professionals who built their careers on the power of a functional curriculum. These educators know how far we have come since the relatively recent days of automatic institutionalization of these students. Some of you may recall seeing film or photos of the hellish conditions of our public institutions in the mid 20th century, and many of you realize that these dreadful conditions were in part what spurred passage of P.L. 94-142. The field of severe disabilities can be very proud of what it has achieved in transforming thinking and actions so that all children are seen as public school students rather than as patients in institutions. Nevertheless, standards-based education has opened the pathway for another huge step for students with significant cognitive disabilities. It is a time of change that is difficult. The field has to carefully examine the evidence to ensure that artificial barriers are not sustained – these barriers include resisting the students’ access to rich and engaging content that has proven to be powerful for their more typical peers. Post school outcomes for students with significant cognitive disabilities based on the functional curriculum as historically defined have been abysmal. It is not a great risk to assume their potential competence in academic content. Indeed there is much to gain.

In many of the states I work with, the current “on loan” teachers who are leading training efforts statewide for the alternate assessments based on alternate achievement standards, but linked to grade-level content, have come from schools – and their own classrooms – where they have taught a functional curriculum for decades. A common story is that these lead trainers once were fiercely opposed to shifts in the curriculum for their students, but were so shocked by what their students could do once they actually gave them the opportunity, that they now help other teachers (and parents) through the initial changes. The student work we are seeing is compelling, and the dialogue of “what is functional” for these students has become extremely convincing. The “least dangerous assumption” appears to be to follow the students’ own evidences of learning, just as many of these gifted teacher trainers have done, and monitor the effects closely over time.

Proposed alternate assessment option. There is a third alternate assessment option to be released in final regulation – alternate assessment based on modified achievement standards, or what some call the “2%” option. I have not been involved in thinking about the development of this option, and cannot at this time speculate what the Rule will say, or how that option will play out in practice. However, going back to the National Research Council Committee on Assessment, it will be critical to understand who the students are for whom the option is designed and how they build competence in the academic domain tests. Comparing this to the established assessments based on alternate achievement standards, we have a pretty good sense of who the “1%” of students with significant cognitive disabilities are, and there is agreement that small group of students can benefit from an alternate achievement standard. The field is still struggling with defining how students with significant cognitive disabilities build competence, or even what the nature of linkage to grade level content can and should be, but the disagreements are honest discussion in pursuit of higher expectations and improved outcomes.

The data from the studies referenced above indicate that the students who are the lowest 2% performing students on regular assessments are a blend of students who may or may not have disabilities, and who predominantly represent student groups who have historically been on the low side of the achievement gap. The data, combined with evidence that many of these students have not been taught the challenging curriculum expected for all students, suggest a need for thoughtful and data-based processes to understand what a modified achievement standard represents. Over the next years, we expect that we will be asked to work with states to mine their data – to understand who the students are and how they build competence – and to support their work in developing tests and defining achievement standards that raise the bar and close the achievement gaps these students represent.

TECHNICALLY SOUND ACCOUNTABILITY SYSTEMS: States should have been developing challenging content and achievement standards and standards-based assessments that include all students as part of Title I requirements since the 1994 IASA legislation. They were required to provide access to the general curriculum for students with disabilities since the 1997 IDEA amendments. Although these laws required public reporting of student achievement for school reform purposes, there really were not public system accountability requirements in either law. Given the track record of progress in most states on raising achievement during this time period, this lack of progress suggests that many public school and state leaders had not read the 1994 and 1997 “memos” from Congress requiring a shift to high expectations and high standards for all students. The NCLB focus on public accountability did catch their attention, fortunately.

Despite the pressure of interesting conversations with constituents that members of Congress must be having, they reflect the awareness of that shift to high expectations and high standards for all students, and yes, for some, a continued disbelief that it is possible. Those of us with a belief that we CAN and MUST do all we can to close the achievement gap, and with data showing that we are moving in that direction, are depending on you to make it possible to continue and finish what has been started. We need to do this by emphasizing good evidence-based practices. We need to do this by reducing tendencies to game the accountability system.

We have all heard the stories of some states and districts that have finessed their accountability plans in the face of bad news from schools. Most states do not have the capacity to “fix” all the schools that could be identified as not doing well with all their students. This may contribute to seeing what some have characterized as a misuse of accountability options, such as setting a very high number for the minimum number of students before results of those students are made public or are included in accountability calculations.

We cannot jump to conclusions about what approaches are “gaming” the system. States have struggled to understand how to avoid over or under-identifying schools for improvement. The technical difficulties of this task are real, and states have an obligation to avoid both false positives and false negatives. We rely on our colleagues with more

technical skills to advise states, and we believe there are good faith efforts in many states to get this right. I have been in conversations in some states where they are designing monitoring checks on schools where the minimum or confidence interval calculations mean that the school was not publicly identified as missing its targets, but the numbers suggest some problems may exist. This is in contrast with other state conversations where the motive appears to be to protect schools at the expense of students. None of us can accurately impute motive to others, but the issue here is that thoughtful, committed people are struggling with ensuring fairness all the way around, and sometimes the rhetoric around accountability means it has been hard to discern who the good guys are from the not so good.

We know we are at a point where there is tension and there is talk about the need for some adjustments in NCLB. Growth models are seen as a logical solution by many, even though like all technical approaches, there are assumptions that should be examined as the models are developed and implemented. Pilots of the models are underway, which is good, and serious attention now has been given to ensuring that all student groups are included, which is better. The states working on this thus far are in pilot phases, and are required to carefully analyze the effects of these models. They also are required to build these models based on an absolute standard of proficiency for all students.

However, many special educators and the general public have seen the term “growth” as more generic – that any progress is acceptable, and would relieve the pressure of “proficiency” as an absolute standard. Some have recommended looking at alternative ways of holding schools accountable for students with disabilities. One example comes from those who assume students on IEPs are already getting high quality instruction on the challenging content, or in some cases, on what they believe should be a separate curriculum. They propose that individualized growth set by an IEP is an example of an accountability system, and could replace the “regular” accountability system for other students. That would have unintended negative consequences for the disability subgroup (and others as well). Students with disabilities for years have shown growth against IEP goals, and we have ample data to show us that has not worked well to raise the bar of achievement (or of expectations) for these students. Others suggest that special education students should be held to separate standards that focus only on basic skills, or should be exempted from accountability completely. This would be a return to where things were before NCLB and before IDEA 1997.

It is important to step back to celebrate where we have come from and to clarify where we cannot go. Because of NCLB and the public clarification and support it has given to high expectations for all students, we now have a powerful lever – perhaps the most powerful one in the past three decades – for reducing and eliminating the achievement gap of students with disabilities. This can happen only if we take advantage of the challenging standards that have been set, and if we apply what we are learning about teaching to them.

We need to focus attention on these learners, along with all other low performing students, not try to hide their performance or get them out of the system. We need to

carefully examine the quality of our current assessments for use in growth models, and carefully examine effects of current growth model pilots on all subgroups. In addition, ANY adjustments to accountability systems should be made for all students, not just one subgroup, with consideration of intended and unintended consequences for students overall and for student subgroups.

The pervasive low expectations for the achievement of students with disabilities in the past and the present must be confronted and addressed. We have ample evidence that students with disabilities can learn the full range of the challenging and interesting curriculum for all children, overcoming years of poor instruction and access to the curriculum. As I said in the opening, the examples we are seeing in formal research and informal identification of successful schools within states demonstrate that the long-term goal of full proficiency is not an illusion. At this point, we can not accept the argument that we should accept far less. The caveat is, of course, that full proficiency IS an illusion if educators continue doing things just as they have been doing them the past 30 years.

Appendix A:

Performance Data Showing Increases for Special Education Subgroup

Figure 1. Percent of Students who Earned a Competency Determination - Class of 2003

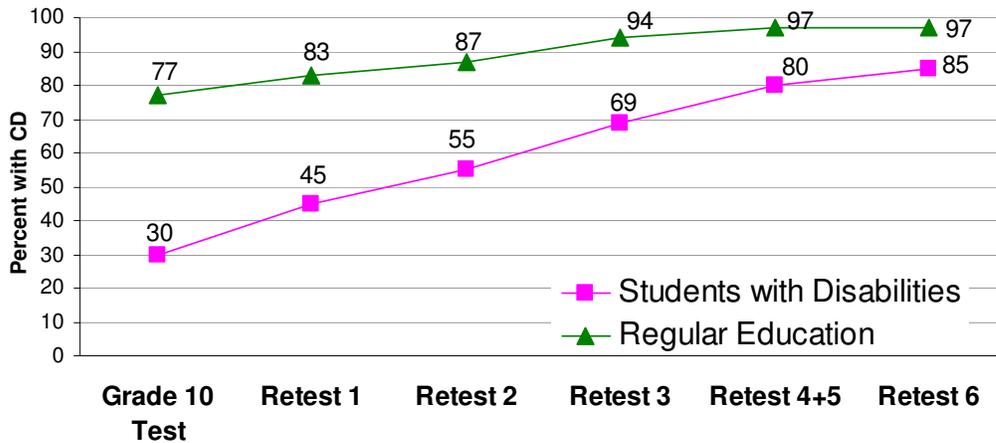
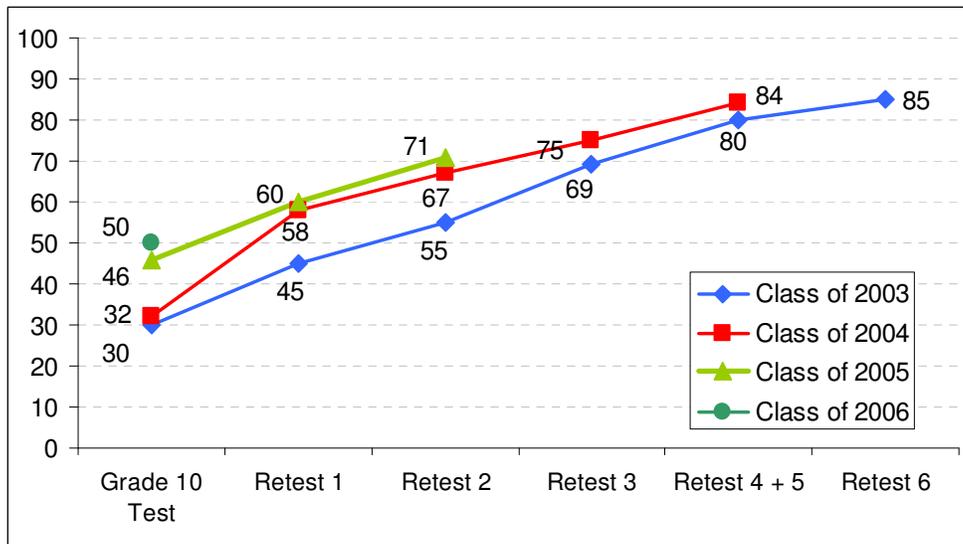
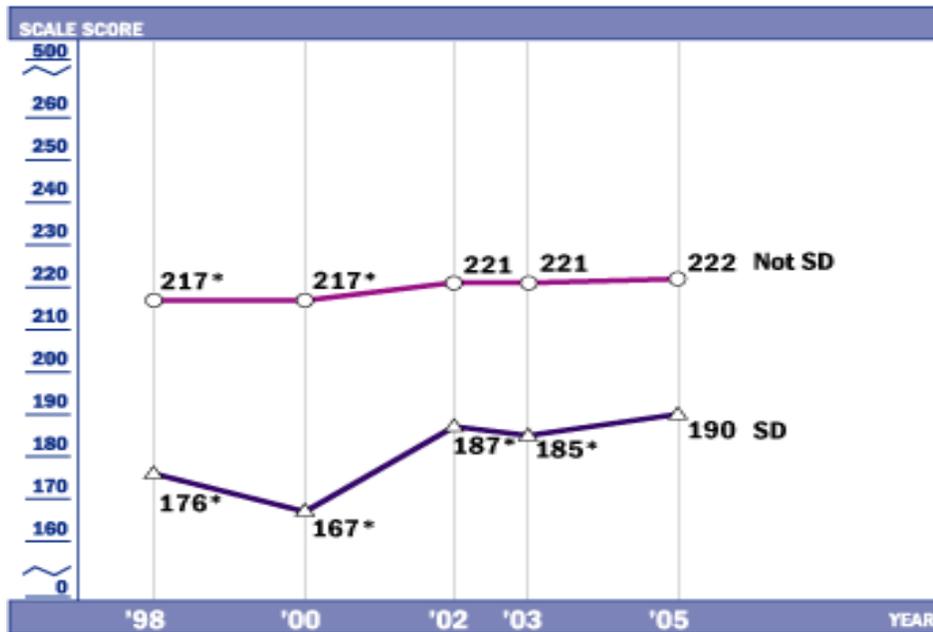


Figure 2. Percent of Massachusetts Special Education Students Attaining the Competency Determination (through the May 2004 MCAS)



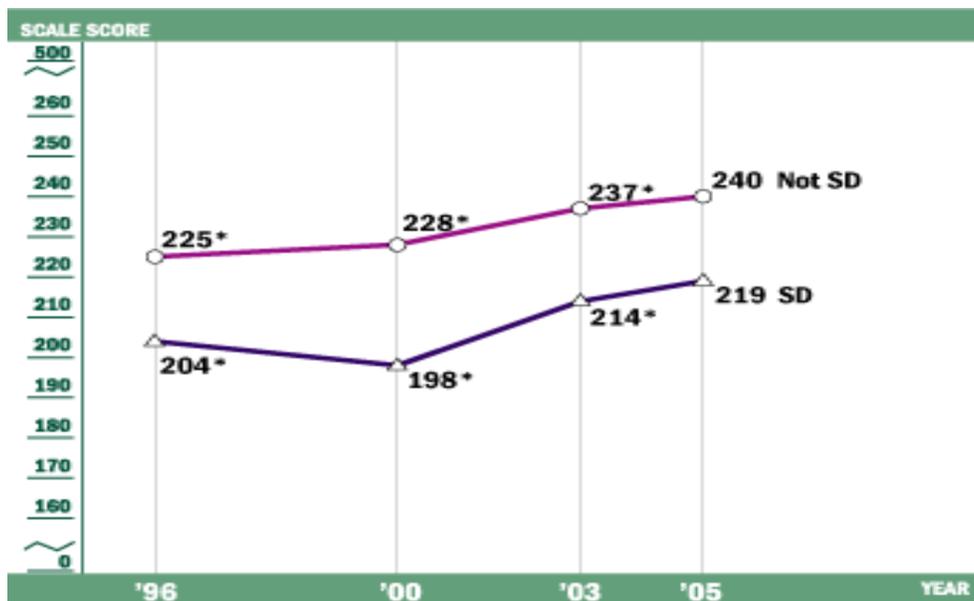
Appendix B:

Figure 3. NAEP Average Scale Scores in Reading, by Students with and Without Disabilities, Grade 4: Various Years, 1998-2005



* Significantly different from 2005. NOTE: SD = students with disabilities.
From: http://nces.ed.gov/nationsreportcard/nrc/reading_math_2005/s0014.asp?printver=

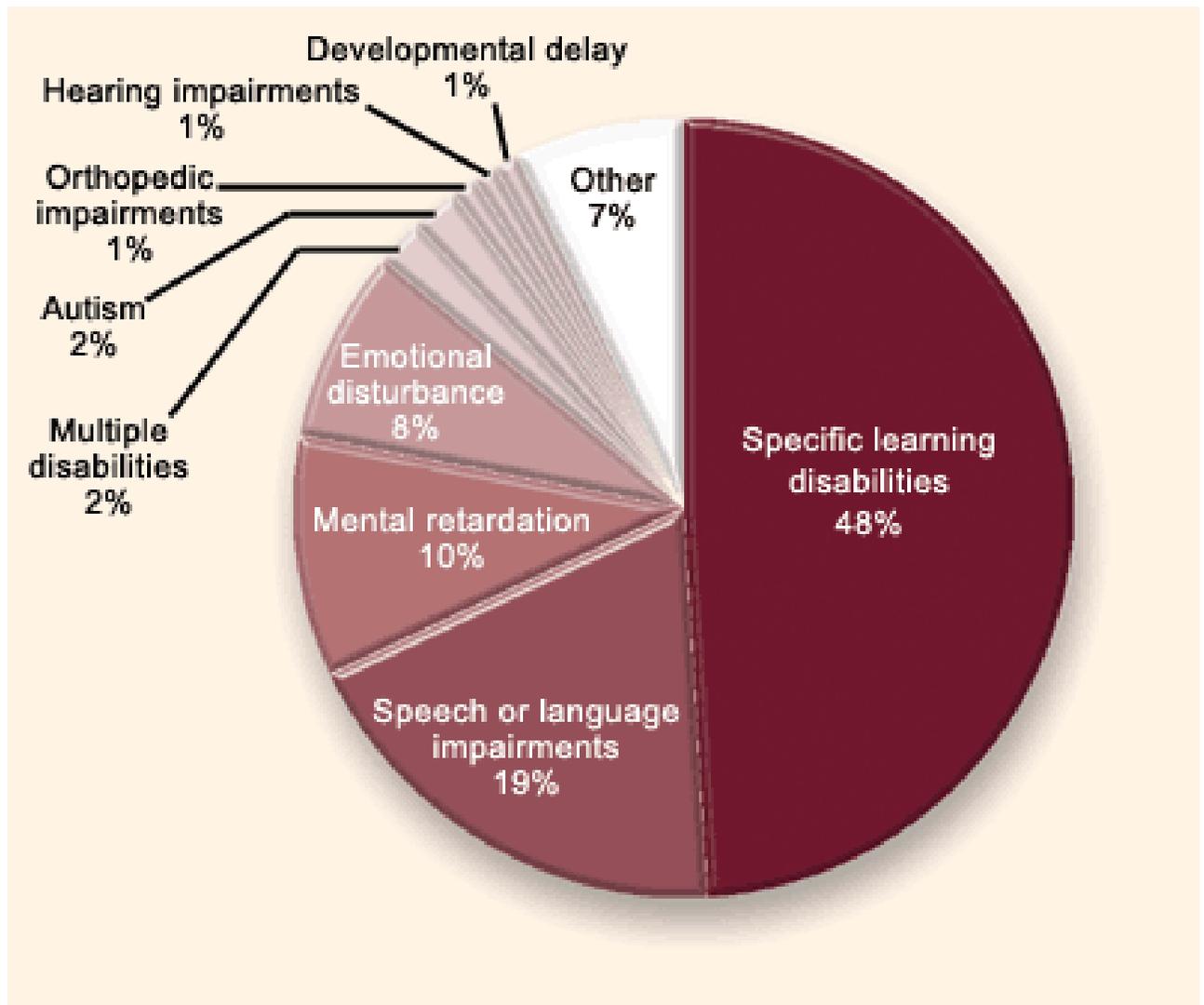
Figure 4. NAEP Average Scale Scores in Mathematics, by Students with and Without Disabilities, Grade 4: Various Years, 1996-2005



* Significantly different from 2005. NOTE: SD = students with disabilities.
From: http://nces.ed.gov/nationsreportcard/nrc/reading_math_2005/s0029.asp?printver=

Appendix C:
Disability Category Population Representation from Education Week Article

In the 2002-03 school year, almost 6 million students ages 6-21 received services under the Individuals with Disabilities Education Act—a number that has increased steadily over the past decade. Those students are classified into 13 different categories under federal law, and the specific needs of each group are very different.



Reproduced from Education Week *Quality Counts 2004, Count Me In: Special Education In An Era of Standards*, p. 10. Information is from an Education Week analysis of data from the U.S. Department of Education, Office of Special Education Programs, Data Analysis System, 2002-03.

References

- Colorado State Department of Education, (2005). *Assessing "Students in the Gap" in Colorado*. Report of the HB 05-1246 Study Committee. Denver, CO.
<http://education.umn.edu/nceo/Teleconferences/tele11/ColoradoStudy.pdf>
- Donahue Institute. (2004, Oct). *A study of MCAS achievement and promising practices in urban special education: Report of field research findings (Case studies and cross-case analysis of promising practices in selected urban public school districts in Massachusetts)*. Hadley, MA: University of Massachusetts, Donahue Institute, Research and Evaluation Group. Available at:
http://www.donahue.umassp.edu/docs/?item_id=12699
- Ferrara, S. (2006). Introduction to the Special Issue Toward a Psychology of Large-Scale Educational Achievement Testing: Some Features and Capabilities. *Educational Measurement: Issues and Practice*, 25(4), pp. 2-5.
- Hawkins, V. J. (2007). Narrowing gaps for special-needs students. *Educational Leadership*, 64 (5), 61-63.
- Horvath, L., Kampher-Bohach, S., Kearns, J.F. (2005). The use of accommodations among students with deafblindness in large-scale assessments systems. Implications for practice and teacher preparation. *Journal of Disability Policy Studies*.
- Kampfer, S., Horvath, L., Kleinert, H., & Kearns, J. (2001). Teachers' perceptions of one state's alternate assessment portfolio program: Implications for practice and teacher preparation. *Exceptional Children*, 67(3), 361-374.
- Kearns, J.F., Kleinert, H. L., Kennedy, S. (1999). Assessment and Accountability Systems: We need not exclude anyone. *Educational Leadership*, 56(6) p. 33-38.
- Lazarus, S.S, Thompson, S.J, & Thurlow, M.L. (2006). *How students access accommodations in assessment and instruction: Results of a survey of special education teachers* (EPRRI Issue Brief Eight). College Park, MD: University of Maryland, Educational Policy Reform Research Institute. Available at
www.eprri.org/products
- National Center on Educational outcomes teleconference presentations:
<http://education.umn.edu/nceo/Teleconferences/tele11/default.html>
<http://education.umn.edu/nceo/Teleconferences/tele12/default.html>
<http://education.umn.edu/nceo/Teleconferences/tele13/default.html>
- New England Compact. (2007). *Reaching students in the gaps: A study of assessment gaps, students, and alternatives*. (Grant CFDA #84.368 of the U.S. Department of Education, Office of Elementary and Secondary Education, awarded to the Rhode Island Department of Education). Newton, MA: education Development Center, Inc.

- Pellegrino, J. W., Chudowsky, N., & Glaser, R. (2001). *Knowing what students know: The science and design of educational assessment*. Washington, DC: National Academy Press.
- Sheinker, J., & Redfield, D. (2001). *CCSSO Handbook for professional development in assessment literacy*. Washington, DC: Council of Chief State School Officers, CAS SCASS.
- Thompson, S. J., Johnstone, C. J., Thurlow, M. L., & Altman, J. R. (2005). *2005 State special education outcomes: Steps forward in a decade of change*. Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.
- Thompson, S.J., Morse, A.B., Sharpe, M., & Hall, S. (2005). *Accommodations manual: How to select, administer, and evaluate use of accommodations for instruction and assessment of students with disabilities*. Washington, DC: Council of Chief State School Officers, ASES SCASS. Also available from OSEP toolkit at – http://www.osepideasthatwork.org/toolkit/accommodations_manual.asp.
- Thompson, S.J., & Thurlow, M.L. (1999). *1999 State special education outcomes: A report on state activities at the end of the century*. Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.
- Thurlow, M.L., Moen, R., & Altman, J. (2006). *Annual performance reports: 2003-2004 state assessment data (Available on the NCEO Web site at <http://www.education.umn.edu/nceo/OnlinePubs/APRsummary2006.pdf>)*. Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.
- Thurlow, M.L., Quenemoen, R.F., Altman, J., & Cuthbert, M. (in press) *Trends in the participation and performance of students with disabilities*. Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.
- Turner, M., Baldwin, L., Kleinert, H.L., & Kearns, J.F. (2000). Consequential Validity of Kentucky's Alternate Portfolio Assessment. *Journal of Special Education*. 34(2), 69-76