

**Testimony of Brian Gong,
Executive Director,
The National Center for the Improvement of Educational Assessment**

**Before the Committee on Education and Labor
Hearing on the “Miller-McKeon Discussion Draft of the
Reauthorization of the Elementary and Secondary Education Act”**

**Washington, D.C.
September 10, 2007**

Chairman Miller, Congressman McKeon, members of the Committee, I am Brian Gong, Executive Director of the National Center for the Improvement of Educational Assessment. I appreciate the opportunity to offer my comments and encouragement to substantially improve *No Child Left Behind* assessment and accountability provisions in the reauthorization of the Elementary and Secondary Education Act.

For the past 20 years—long before *No Child Left Behind*—I have worked on improving assessment and accountability systems to help foster student achievement and school capacity. I know firsthand the positive influence that good assessment and sound accountability systems can have to promote equitable student learning and deep school improvement. I also know the difficulties of doing it right. As a research scientist at Educational Testing Service in the 1980’s I worked on developing innovative instructional assessments that would support classroom learning and teaching. In the mid-1990’s I served as the Associate Commissioner of Curriculum, Assessment, and Accountability in the Kentucky Department of Education. (Kentucky, one of the first and longest-tenured state accountability systems, is notable for tackling the technical challenges of scoring, reliability, and large-scale administration of performance-based, non-multiple choice assessments. Kentucky still uses a writing portfolio in its accountability system.) Our non-profit Center for Assessment is currently working with 20 states across the nation to provide technical assistance in one form or another to support assessment and accountability systems that are educationally and technically sound. The Center for Assessment is also regularly called upon to provide technical assistance in these areas, by groups including the U.S. Department of Education, the Council of Chief State School Officers, and the National Center on Education Outcomes. I recently served on the Expert Panel on Assessment convened by the Forum on Educational Accountability; the final report, *Assessment and Accountability for Improving Schools and Learning* (2007) is available here today and addresses in more depth many issues relevant to reauthorization. The Center for Assessment is also working on areas outside of NCLB, including formative assessment and college readiness with some states and organizations including Achieve and the Gates Foundation.

My comments fall in two main areas:

- I applaud the recognition for some mid-course corrections to ESEA. Several of the provisions of the discussion draft respond to concerns, but need some tuning in the legislative solutions.

- I strongly support provisions in the discussion draft that move from fixing “what is” to pointing us where we need to go in the future of assessment and accountability.

Some Mid-Course Corrections

I comment on several areas in the discussion draft that courageously acknowledge some problems in the 2001 legislation and undertake making mid-course corrections.

1. To be more valid, school accountability should be broadened to include student growth. Everyone is concerned about whether schools helped students *learn* during the year, not just how high they scored. Unfortunately, our recent studies show that the way that growth has been defined by the U.S. Department of Education in its Growth Model Pilot program actually hardly differs from Status (percent proficient). Reauthorization should include a true pilot of how growth could effectively be measured and used for accountability.
2. It absolutely makes sense to distinguish between a school that fails to meet a few of the hurdles from a school that fails to meet many. But, there are many cases where performance of one or two subgroups are not only very important but can be measured very reliably. Reauthorization should consider ways to make meaningful, reliable distinctions besides just counting the numbers of students and subgroups and making a decision each year.
3. “Minimum-n” subgroup sizes and especially confidence intervals are important safeguards to support making reliable accountability decisions. Setting common thresholds across states makes sense from a technical standpoint. Hopefully the final reauthorization version will alter slightly the thresholds. Our studies show that a 95% confidence interval *for the overall decision*—rather than for each subgroup/content area decision—would help avoid as many as 20% of the schools being misidentified.
4. In my opinion, the aspirational goal of 100% of the students proficient by 2013-14 is not a credible goal. It is possible to define goals that will be challenging, rigorous, equitable, and possible. In 2004-05 over half of the states had already identified at least 20% of their schools as not meeting AYP; seven states had identified over half of their schools. Even more will be identified next year when the AMO targets are increased. The reauthorization must address this fundamental issue.
5. Working with states on accountability issues over the past 15 years, I have become more convinced that strong accountability systems are important, and can be helpful, but are not enough. In fact, many schools do not know what to do to improve, and many face serious structural barriers, such as hiring and retaining strong teachers who are effective with the students in the school. The reauthorization and any school improvement plan must have a better theory of action than saying “Clear goals and strong sanctions will motivate schools and districts to solve this problem.” I simply do not believe that is true; it is not a helpful characterization of the problem or the solution to improving American education. The discussion drafts attention to improved professional development, coupled with an improved accountability system is a step in the right direction, but needs to go much further in terms of strong models of support.
6. We need to work to include accountability special populations in meaningful ways. However, our current attempts at assessment of students with disabilities and students with limited English proficiency reflect more noble policy aspirations than sound measurement. Reauthorization should take a more realistic look at what is scientifically

known about good assessment and learning, and inform the accountability requirements accordingly.

7. Much of the complaints from the states reflect not so much the statute, but the process of interacting with the U.S. Department of Education. Reauthorization would do well to attend to how the process of interpreting, enforcing, and supporting the implementation of the law is done, not only by the states, but also by the federal Executive Branch.
8. I think that content standards, assessments, and accountability must be yoked together with equally strong curriculum and instruction in order to have effective learning and teaching. I do not believe that movement towards federal or national standards can be effective without equal attention to curriculum. Reauthorization must pay attention to the debate of the proper role of the federal government in establishing supra-state standards.

Support for Draft's Vision of Investing in Future Assessment

I strongly support provisions in the discussion draft that include support for a wider and more valid set of assessments, including performance assessments. This is not about going soft on accountability. This is about creating incentives to develop assessments that validly reflect what the next generations of American students truly need to know and be able to do. The proposed legislation is a good step in that direction.

Some people may portray this as a backdoor attempt to water down accountability or to undermine rigorous standards. I don't read the discussion draft that way, and I wouldn't support it if I thought it did. I read the draft as providing incentives to try to develop more advanced assessments, including performance-based assessments; it provides a clear mandate that such assessments are not to be used for accountability unless and until they meet rigorous criteria administered by the U.S. Department of Education.

I support the discussion draft's attention to three longer-term needs in assessment.

1. It helps us attend to some very important skills that are simply not possible to assess well in current traditional assessments, particularly several aspects associated with college and work readiness.

Problems of college readiness will not be solved largely by having a more stringent graduation standard, a longer end-of-course exam in Algebra, or federal performance standards for what it means to be proficient. Certainly we need to ensure that high school students have the academic knowledge these things represent. But success in college and success in life requires a whole set of additional skills than are currently being assessed. These skills have been called "habits of mind" by some. That's a fancy title, but the skills are familiar—extended problem solving, ability to do research, write clearly, monitor one's own performance to be sure it is appropriately accurate and precise. In addition, we've heard for decades that employers care more about what graduates can *do* than what they *know*, and even more about how they actually perform in real world settings, not the artificial confines of a standardized test. That's why employers and colleges are both interested in performance assessment and documentation of such things as ability to communicate well orally, ability to work well in a small group, honesty, self-discipline, responsibility for getting the work done.

I believe that we all would agree that such things are important and that such things are not being assessed at all in our current assessments. I believe that we can do much more to assess such vital college readiness skills, and do it in a way that is valid, reliable, affordable, credible, and useful. The alternative is to do nothing. And then, even if the grand goal of NCLB is reached in 2013-14, we'll find that we have students who can spit back answers on a multiple-choice test, with perhaps a few short answers, and even perhaps solve some pretty hard Algebra items about polynomial functions—but they may not be any better prepared to succeed in college, work, or life. The discussion draft represents an attempt to seize this opportunity to invest even a modest amount in assessing those essential learning skills that really matters, which we're not doing now.

2. It provides an **investment in the future infrastructure of assessment**, such as complex performance assessments, the use of technology, and advanced psychometric models that incorporate what is known about how people learn.

It is true that there are current technical and operational challenges to using performance assessments at large-scale for high stakes purposes. The road for implementing complex assessments has been rocky. That is exactly why the field and the nation need the investment outlined in the discussion draft. For example, our children already play computer games that immerse them in realistic role-playing simulations, distributed group competitive strategies, and that support voice and motion recognition. I cannot but imagine that in 20 years computers will have even more capacity. But I can imagine that unless an investment is made, educational testing in 20 years will be as hobbled by a lack of imagination and by 19th century measurement theories as it is today. Reauthorization should look to the future as well as try to make mid-course corrections to the present.

3. It provides needed **federal sponsorship** that will catalyze partnerships and applications that will address and sustain the effort to develop new assessment infrastructure.

I believe that universities, research centers, and the private sector together can help bring the next generation of valid assessments to the schools. But it won't happen without a catalyst to focus the use, practicality, and time schedule. The federal government can appropriately provide that sponsorship, as is proposed in the discussion draft.

It's a good time for mid-course corrections, and for investing in the future of college-ready and performance assessments. I urge Congress to support the suggestions for reauthorization I've mentioned today.

Thank you for the opportunity to share these thoughts.