

Testimony on  
"ESEA Reauthorization: Options for Improving NCLB's Measures of Progress,"  
before the Committee on Education and Labor  
U.S. House of Representatives

Harold C. Doran  
American Institutes for Research  
Senior Research Scientist

March 21, 2007

Chairman Miller, ranking member McKeon, and honorable members of the committee, thank you for this opportunity to share my thoughts on ways to improve the No Child Left Behind Act. My name is Harold Doran, and I am a senior research scientist at the American Institutes for Research (AIR) in Washington, DC. In this role, I help states and districts across the country develop their testing and accountability systems.

The question I have been asked to respond to is whether the adequate yearly progress (AYP) provisions in NCLB would benefit from having additional ways to evaluate schools, what some refer to as *multiple measures*, and whether these measures can be joined to form a compensatory accountability system. The term *compensatory* denotes that not meeting AYP under one measure could be compensated for using a secondary measure.

I believe the AYP provisions could be strengthened if multiple measures were added. In my discussion today, I would like to explain this position and suggest specific measures that I believe would strengthen the legislation.

### **Why Multiple Measures?**

I emphatically support the use of multiple measures, as do most educational experts. However, there are multiple views on what set of measures to include in accountability systems. Even more challenging is how these measures can be combined in forming a compensatory accountability design. To reduce ambiguity, I would offer the following definition of multiple measures for today's conversation: An accountability system that includes multiple measures uses test scores from more than a single test, achievement indicators collected by other means, or various statistical methods for evaluating the data.

By this definition, NCLB already relies on multiple measures. But the law does not permit one measure to compensate for another measure. I believe the integrity and strength of the law would be enhanced if it were modified to accommodate the following:

1. Permit for multiple measures; and

2. Allow states to use those measures to create rigorous compensatory systems.

Any consideration of new measures, however, must first be met with a discussion of criteria to avoid watering down our current systems:

1. **Increased Rigor.** Including new indicators should result only in added rigor to core content areas.
2. **Simplicity and Transparency.** Incorporating multiple measures should not result in complex systems that are difficult to implement or that are confusing to parents and educators. The elegance of simplicity, combined with a focus on rigor, will guard against over-engineering accountability designs.

### **Specific Recommendations for Multiple Measures**

I have four specific recommendations. Two of these recommendations would add measures that could serve in a compensatory role, one recommendation adds to AYP, and the last is a recommendation to ensure system integrity.

### **End-of-Course Exams**

NCLB currently monitors the proficiency rates of high-school students in language arts/reading and math. When students do not reach levels of proficiency on the statewide regular tests, their only option in many cases is to retake the same test. However, an alternative that could be used is to provide students with an opportunity to enroll in coursework that targets their specific areas of need and allow for them to pass an end-of-course test that demonstrates mastery of the content.

For instance, a student may not reach proficiency on the statewide NCLB test only because he struggles with concepts in geometry. Subsequently, the student could enroll in a geometry course and, at the end of this course, demonstrate proficiency via a state-developed end-of-course exam in geometry that is equally as rigorous as the statewide NCLB test.

### **Growth Models**

Learning is fundamentally about change. However, the methods by which AYP are currently calculated do not follow this logic and are, in many ways, biased.

The current reality is that the mathematical model used to measure proficiency rates must be improved. For example, a school with many students scoring in the highest performance category can have a drop in students' academic performance that still remains above the proficiency bar and still be classified as

making AYP. In contrast, a school with many students beginning well below proficiency, but learning at remarkable rates, is likely not to be recognized as a high-performing school.

It is my recommendation that AYP calculations include results obtained from growth models as another method for evaluating schools. The results from these models can be used in a manner similar to the safe-harbor provisions as another way to make AYP. If permitted, the models must conform to the same high expectations for proficiency as currently required and not simply reward growth.

### **Incorporate Science Results into AYP**

The 2001 NCLB requires students to participate in science assessments beginning in 2008. However, the results of those science assessments are not included in the current AYP calculations.

Including science in AYP calculations will encourage schools to emphasize science as a component of their core curricula. It will also be possible to develop end-of-course exams in science as previously suggested.

### **National Assessment of Education Progress (NAEP) Research for Comparability**

Last, I would like to offer a suggestion on the use of NAEP—it cannot be used to measure AYP, but it can be used to inform how state performance standards are set and partly used to determine overall system integrity. I would like to recommend that this committee support a research agenda that would investigate and report how best to establish links between NAEP and the various state assessment programs across the country.

In many respects, the variability in content standards and difficulty of the assessments across states is important and reflects critical idiosyncrasies in the educational programs. On the other hand, this variability presents a significant challenge given that we live in a highly mobile society. For example, a student attaining mathematical proficiency in Arizona may attend college and/or obtain professional work outside of that state.

Hence, my view is that reauthorized versions of NCLB should establish national policy using NAEP to illustrate the comparability of proficiency levels across the country. This information would be extremely valuable as states build and/or refine their standards and assessment programs. It will also provide policymakers with a window to assess system integrity.

Should the committee accept the notion that additional indicators are necessary to establish more robust systems, I would then encourage the committee to further consider how these multiple indicators can be combined to form a

judgment about school quality that still aligns with the basic tenets of proficiency set forth in the legislation.

I hope these suggestions are helpful as this committee moves forward with deliberations related to NCLB improvements. I am grateful for the opportunity to testify today and am happy to answer any questions you may have.