

Comments on Selected Provisions in the Every Student Succeeds Act

Stuart Kahl | December 2015

As of this writing, the bipartisan bill for the reauthorization of the ESEA is well on its way to becoming law—possibly before the end of this month. Although eight years late in coming, the Every Student Succeeds Act has retained many of the provisions of NCLB, but has also included some new or changed ones. Below I discuss a few provisions that I believe have important implications regarding accountability assessment. There are many others.

“Annual” Testing

Actually, this term really means annual testing in grades 3 through 8 and one high school grade in English language arts and mathematics, as opposed to annual testing in fewer grades as in the early years of NCLB. This is not a new provision, but concerns about excessive testing led to efforts by some to reduce the amount of testing by going back to the earlier approach. In the end, the requirement for testing in seven grades in the two subjects remains, and I believe it is worthwhile to mention three good reasons for its continuance.

- The monitoring of student growth over time is enabled by the testing at consecutive grades.
- The “annual” testing requirement is supported by various groups advocating for students with special needs (students with disabilities, English language learners, etc.). These student groups have benefitted significantly from the attention NCLB has given them.
- Something many seem to have forgotten, when testing occurred at just a few non-contiguous grades, the measurement community had to

warn schools that their results could fluctuate a good bit from year to year simply because of ability differences of the groups of students passing through the tested grades. Their solution was to not put too much stock in one year’s results and to aggregate multiple-years’ results before making important decisions. Testing at all the grades from 3 through 8 was another solution to the problem, a solution that allowed more timely decision making.

“In the end, the requirement for testing in seven grades in the two subjects remains, and I believe it is worthwhile to mention three good reasons for its continuance.”

Multiple Measures

According to NCLB, assessments were to “involve multiple up-to-date measures of academic achievement, including measures that assess higher-order thinking skills and knowledge.” Little attention was given to this requirement in many state assessment programs, primarily because the reporting requirements and increased amount of

testing necessitated the use of time- and cost-efficient testing approaches. Thus, the term “multiple measures” came to mean “multiple-choice items with token use of items of some other format.” Interestingly, an early version of the Senate bill interpreted the term to mean “status and growth measures.” This hardly captured two important ideas underlying the call for “multiple measures”—the basic “rule” about not basing important decisions on a single measure, and the reality that many accountability assessments had been short-changing the measurement of higher-order cognitive skills.

“The flexibility regarding multiple measures and through-course testing, should it remain, would allow for considerable creativity in the design of state assessment programs.”

While the current bill, soon to be passed, accommodates the measurement of both status and growth, it also requires the measurement of higher-order thinking skills. It goes so far as to say that the assessment may be “partially delivered in the form of portfolios, projects, or extended performance tasks.” The bill also allows single summative assessments or multiple assessments during the course of the year, the results of which would be combined to provide reliable and valid summative scores.

This flexibility regarding multiple measures and through-course testing, should it remain, would allow for considerable creativity in the design of state assessment programs. For example, I’d like to see

a program with very brief end-of-year summative components combined with state-approved, curriculum-embedded performance assessments used at several times during the school year.

The bill also allows a limited number of states to try out locally developed assessments in place of state assessments. However, I believe it is unlikely that locally designed assessments, without state involvement in their development and piloting, would meet the requirements for technical quality, in particular the requirement for comparability of results across all schools in a state.

Reporting

Both bills borrow from the original NCLB language regarding reporting. Of course, the requirement for disaggregated results by various subgroups is a hallmark of the legislation, consistent with its requirement to include “all students” in the assessments. Making schools accountable for the performance of all their students was something new and gave special populations of students the same status as the “general” population of students.

Another reporting requirement is not new, but merits comment—the NCLB requirement for student-level reporting in both bills. The assessments must “produce individual student interpretive, descriptive, diagnostic reports.” The idea that a short, summative assessment sampling a whole year’s worth of content in a subject area/domain could be particularly diagnostic is simply ludicrous. Fortunately, a few subtest scores and general descriptions of students’ capabilities at each of four performance levels seem to have satisfied this requirement in the past. Nevertheless, reporting requirements often, more than anything else, dictate the design of an assessment program. In the future, if a state should choose to combine a brief summative component with curriculum-embedded performance assessment, the latter component could be the basis for far more meaningful individual student reporting.

Accountability Requirements

Supporters of the original NCLB most certainly considered the accountability requirements in that law to be the “teeth” of the legislation—the real force behind the desired improvements in student achievement. Yet in addition to expressly forbidding the Secretary from dictating the content standards and assessments the states are to use, the current bipartisan bill also leaves it to each state to determine how it will use the assessment results for accountability purposes.

These changes in the bill may not have occurred if the USDOE had not placed three requirements on states in order for them to receive Race to the Top monies:

- adoption of the Common Core State Standards,
- membership in one of the two major state assessment consortia, and
- significantly weighing student assessment results in the evaluation of teachers.

The backlash against these requirements has ultimately led to the bill’s inclusion of specific restrictions on federal authority (already intended by the Constitution, which left education a state responsibility).

Overall, with respect to accountability assessment, I believe the Every Student Succeeds Act, relative to its predecessor NCLB, provides greater clarity, offers greater flexibility to states, and reduces the likelihood of misuse of test results.