

# **Forum on Educational Accountability**

<http://www.edaccountability.org>

## **Overhauling Assessment Systems**

No Child Left Behind requires that every state test every child annually in reading and math in grades 3-8 and once in grade 10-12. These tests are the primary, near-sole basis for evaluating schools and districts under NCLB. Under Race to the Top and the Department of Education waiver process, states must use student test results as a “significant factor” in judging teachers [and principals?]; more than half the states are now doing so, including some that have not been awarded RTTT or a waiver. The requirement to evaluate teachers also means that either new tests must be devised to test students in untested grades and subjects, or to use current tests (e.g., math and reading) to judge teachers in other subjects (e.g., science, history, art, phys ed, second language). These federal policies have led to a vast expansion in the quantity of standardized testing as well as in their high-stakes uses.

### ***Joint Statement on NCLB***

The founding document of the Forum on Educational Accountability -- the *Joint Organizational Statement on No Child Left Behind*, which has been endorsed by more than 150 national education, civil rights, religious, disability, parent and civic organizations – proposes a different approach to assessment, as well as accountability, improvement and equity.<sup>i</sup> It states that the federal government should support development of assessment systems that:

- Provide a comprehensive picture of students' and schools' performance by moving from an overwhelming reliance on standardized tests to using multiple indicators of student achievement in addition to these tests.
- Help states develop assessment systems that include district and school-based measures in order to provide better, more timely information about student learning.
- Strengthen enforcement of NCLB provisions requiring that assessments must:
  - Provide multiple, up-to-date measures of student performance including measures that assess higher order thinking skills and understanding; and
  - Provide useful diagnostic information to improve teaching and learning.
- Decrease the testing burden on states, schools and districts by allowing states to assess students annually in selected grades in elementary, middle schools, and high schools.

### ***FEA Summary Recommendations on Student Assessment***

The implementation of these recommendations will require major changes in assessment systems. FEA summarized core recommendations to the Congress for overhauling federal law on assessment, and informed Congress of its views on testing and educator evaluation.<sup>ii</sup>

Regarding student assessment, FEA recommends<sup>iii</sup> that Congress and the Administration:

- Reduce the amount of mandated testing (e.g., return to requirements in the 1994 federal law of once each in elementary, middle and high schools), thus aligning the U.S. with the practices of most nations in which fewer but better assessments produce superior results.
- Support development of state and local assessment systems that include classroom-based evidence as part of public reporting and accountability, and for improving teaching and learning.
- Support development of improved assessments and assessment elements, such as performance tasks and projects, which states can make available to educators to use when appropriate and incorporate into large-scale assessments.
- Require states to use multiple sources of evidence of various types (“multiple measures”) in evaluating students, schools and educators, and in constructing any growth/improvement/value-added approaches.
- Ensure that new assessments consider the needs of diverse learners, including use of the principles of universal design for learning.

FEA also recommends that these principles apply to the Department-funded development of Common Core State Standards-based assessments developed by multi-state consortia.

### ***Expert Panel on Assessment Report***

The summary recommendations to Congress are based on the *Joint Statement* as elaborated in the report of the Expert Panel on Assessment, *Assessment and Accountability for Improving Schools and Learning: Principles and Recommendations for Federal Law and State and Local Systems*. That report explained how to develop a system that would meet the FEA goals expressed in the *Joint Statement*. In essence, the Panel outlined a system built from the classroom out. It takes the form of principles and recommendations. Below are key components of such a system, excerpted from the Panel report.<sup>iv</sup>

#### *Principle I: Equity and Capacity Building for Student Learning*

Help states, districts, and schools fulfill their educational responsibilities to foster student learning and development by ensuring that all students have equitable access to the resources, tools, and information they need to succeed and by building capacity to improve teaching and learning.

#### *Principle II: Comprehensive State and Local Assessment Systems*

Construct comprehensive and coherent systems of state and local assessments of student learning that work together to support instruction, educational improvement and accountability.

1. Provide incentives for states and districts to develop comprehensive and coherent assessment systems that inform instruction and decision-making in ways that state tests alone cannot and do not. Coherent and comprehensive assessment systems provide evidence of student and school performance in relation to rich and challenging educational goals, using multiple indicators of student learning from a variety of sources at multiple points in time.

This would provide resources to enable development of assessments at all levels, particularly addressing important standards that cannot be assessed well with current large-scale assessments, such as reasoning, communication, problem-solving, research, oral communication, and applied learning. Comprehensive

assessment systems would address these areas through employing multiple appropriate assessment practices and tools, including: teacher observations; tests that include multiple-choice, short and longer constructed response items; essays; tasks and projects; laboratory work; presentations; and portfolios. It would also include development of assessments for specific subgroups, including English language learners (ELLs) and students with disabilities (SWDs).

2. Provide states incentives and supports to include high quality local assessment systems in meeting ESEA's accountability requirements, alone or by augmenting state assessments. Fund pilot projects in which interested states demonstrate how they can meet ESEA's accountability requirements through standards-based, locally-developed assessments of students' learning or by integrating local assessments with state assessments. Fund expansion of the number of supported projects as states indicate interest. Provide incentives for states to work together.
3. Provide tools for states and districts to self-evaluate and improve the coherence and effectiveness of their local comprehensive assessment systems. The assessment and instructional components should work together to support instructional improvement and educational accountability.
4. Design state and district reports that include all components of their comprehensive assessment system. State reporting should reflect the presence and weighting of the results from each measure that comprises its assessment system, including those unique for ELLs and SWDs. District level reporting should mirror the results of measures, both local and state, used in its accountability system.

In discussing these improvements to the law, the Panel explained:

Local assessments include teacher-based assessments (their own or ones adapted from other sources) conducted in their own classrooms. They also include “common” assessments, which are developed or adopted by groups of teachers at the school and district levels. Local assessments may include specific instruments or practices (e.g., tests and projects, teacher observations or discussions with students) and compilations of evidence of learning (e.g., portfolios, exhibitions and learning records).

Local assessments perform two legitimate, important, different, and complementary roles. First, they provide information to guide teaching and local and classroom instructional decision-making that state tests cannot and do not provide, and as such, they can operate parallel to the state assessment system without actually being integrated into it.

For example, “formative” assessments – ongoing assessments used to shape instruction and provide feedback to students — have been shown to be powerful tools to assist student learning and provide greater assistance to low-performing students. The use of such formative assessments must be primarily in the hands of teachers. Not all “summative” assessments (i.e., those that provide periodic summary evaluations of student learning such as at the end of a semester or year) are or should be designed for use in accountability; much summative evidence is properly the domain of teachers, students and schools.

A second, equally appropriate use of local assessments is to build some of them into the state's accountability system to help states meet ESEA's accountability mandates. They can provide locally-based valid and reliable summative information about students' learning. The use of locally developed assessments as acknowledged components of the ESEA

accountability system would assign to states the responsibility to assist local districts in developing assessments with sufficient fairness, reliability and validity to meet established accountability requirements.

Questions of how to assemble and weight multiple components from different sources could be addressed in the state pilot programs we recommend. However, no one test should be weighted so heavily that educators effectively ignore other assessments because the one test is what really “counts.”

### *Principle III: Assessment and Accountability for Diverse Populations*

Shape the design, construction, and application of assessment systems so they are appropriate for an increasingly diverse student population.

1. Design assessments based on principles of universal design, but ensure that the unique factors that impact the performance of subgroups (e.g., English language learners (ELLs), students with disabilities (SWDs), students from major racial and ethnic groups, or economically disadvantaged students) are specifically addressed in the assessments that are used to measure the academic achievement of these students and reporting of results.
2. Require states to provide research-based recommendations for selecting and using appropriate accommodations for ELLs and SWDs to ensure that these students have access to valid assessments of their content knowledge.
3. Require states to validate assessment systems for each subgroup.
4. Support research to address major issues that complicate the design of appropriate assessment systems for subgroups.
5. Provide incentives for states to work together to shape the conceptual design and construction of local and state assessments of academic achievement according to the characteristics of each specified subgroup. Federally fund research to address the most pressing technical issues related to assessments and accountability decisions for English language learners and students with disabilities.

### *Principle IV: Fair Appraisal of Academic Performance*

Use multiple sources of evidence to describe and interpret school and district performance fairly, based on a balance of progress toward and success in meeting student academic learning targets.

1. Encourage states and districts to use multiple sources of evidence drawn from their comprehensive and coherent systems of classroom-, school- and district-based assessments to summarize and appraise student performance.
2. Encourage states to describe school performance in terms of status, improvement, and growth, using the states' multiple sources of evidence.
3. As states evaluate their assessment systems, conduct ongoing studies of the validity of the descriptions and interpretations of student and school performance to ensure the quality of core data analysis and reporting.

### *Principle V: Fair Accountability Decisions*

Improve the validity and reliability of criteria used to classify the performance of schools and districts to ensure fair evaluations and to minimize bias in accountability decisions.

*Principle VI: Use of Assessment and Accountability Information to Improve Schools and Student Learning*

Provide effective, targeted assistance to schools correctly identified as needing assistance.

1. Encourage states and districts to use multiple sources of evidence from state and local assessments and other forms of evidence to inform actions such as interventions and technical assistance.

***Educator Evaluation***

The Forum on Educational Accountability (FEA) believes the most important purpose of educator evaluation is to improve teaching, learning and schools.<sup>v</sup> Substantial research concludes that current student tests cannot reliably, validly and fairly be used to judge educators. The negative consequences of NCLB predictably will intensify if educators are required to be judged “in significant part” by student test scores. FEA makes the following specific recommendations:

- 1) Congress should not mandate that states construct statewide evaluation systems as a condition of receiving funds under ESEA.
- 2) If Congress creates a voluntary grant program, it should not mandate any particular, fixed weighting for student learning in general or for test scores in particular.
- 3) Congress should require that educators and local leaders play a major role in the design and implementation of any federally-funded state or local educator evaluation program. Parents, students and various civic and other organizations should be included.
- 4) Congress should require the independent review of federally funded educator evaluation programs for effectiveness and beneficial or harmful consequences.

***Bibliography***

- Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2004). *Assessment for learning – putting it into practice*. Buckingham: Open University Press.
- Brookhart, D.M. (2003). Developing measurement theory for classroom assessment: Purposes and uses. *Educational Measurement: Issues and Practice* 23 (June): 17-25.
- Commission on Instructionally Supportive Assessment (2001). *Building tests to support instruction and accountability: A guide for policymakers*. Report commissioned by NEA, NAESP, NASSP, NMSP, & AASA. Available at: <http://www.nea.org/accountability/buildingtests.html>.
- Gallagher, C. (2007). *Reclaiming assessment: A better alternative to the accountability agenda*. Heinemann, Portsmouth, New Hampshire.
- Learning Record. Describes use of this classroom-based assessment process that has also demonstrated validity and reliability for larger-scale reporting. <http://fairtest.org/learning-record>.
- New York Performance Standards Consortium. Describes a performance assessment system in use as dozens of NY high schools. <http://performanceassessment.org/>.
- Pellegrino, J. W., Chudowsky, N., & Glaser, R. (Eds.) (2001). *Knowing what students know: The science and design of educational assessment*. Washington, DC: National Academy Press.
- Position Paper on Assessment for Learning from the Third International Conference on Assessment for Learning. <http://fairtest.org/sites/default/files/Assess-for-Learning-position-paper.pdf>.

National Research Council (2003). *Assessment in support of learning and instruction: Bridging the gap between large-scale and classroom assessment*. Committee on Assessment in Support of Learning and Instruction, Board on Testing and Assessment, Center for Education, Division of Behavioral and Social Sciences and Education. Washington, DC: National Academy Press.

Wilson, M., & Bertenthal, M. (2005). *Systems for state science assessment*. Washington, DC: National Academy Press.

---

<sup>i</sup> The full statement with list of signers is available at [http://www.edaccountability.org/Joint\\_Statement.html](http://www.edaccountability.org/Joint_Statement.html).

<sup>ii</sup> FEA documents, when not signed by specific organizations, represent a consensus approach within FEA that may not represent the precise views of all FEA member organizations.

<sup>iii</sup> The full FEA summary recommendations are available at [http://www.edaccountability.org/FEA\\_2\\_Page\\_Summary\\_Recommendations\\_2011\\_final.pdf](http://www.edaccountability.org/FEA_2_Page_Summary_Recommendations_2011_final.pdf).

<sup>iv</sup> This paper mostly deletes the recommendations under Principles I, V and VI, as they address the issues of resources and then accountability uses of assessments, rather than assessments themselves. It pastes in limited excerpts from the explanatory discussion of the recommendations. The full paper is available at <http://www.edaccountability.org/AssessmentFullReportJUNE07.pdf>.

<sup>v</sup> Excerpted from “What Should Congress Do about Teacher Evaluation?” available at [http://www.edaccountability.org/What\\_Should\\_Congress\\_Do\\_About\\_Teacher\\_Evaluation- FEA\\_letter\\_9-20-11.pdf](http://www.edaccountability.org/What_Should_Congress_Do_About_Teacher_Evaluation- FEA_letter_9-20-11.pdf).