



SANDLER
FOUNDATION

What Multiple Measures Are States Using?

With the advent of the No Child Left Behind Act, most states focused their accountability systems on the federally defined measures of mathematics and reading standardized test outcomes. But at least 20 states maintained independent accountability systems and published school rankings on the basis of their state system as of the 2008-2009 or 2009-2010 school years.

Those states publish information, including additional indicators that they believe are relevant for informing the public, implementing accountability measures, and developing a more sophisticated understanding of the contexts in which their schools operate. Indicators are used for a variety of purposes. Some are used for diagnostic and planning purposes, while others are used for accountability.

➤ **College readiness measures were the second-most common indicator in state accountability systems after test scores.**

- Nine of the 14 states that provided college readiness measures attached stakes to them; the other five used them strictly for informational purposes. The measures included:
 - participation rates and average scores on the SAT, ACT, and Advanced Placement tests
 - participation in advanced coursework, including dual enrollment college courses
 - remediation rates for high school graduates upon entering college
 - dropout and graduation rates by students who are “at risk” were made weighted measures in some states to create an incentive for schools to keep those students in school
- Some (rural) states also applied weights to the type of high school diploma received (e.g. technical/vocational, GED, or standard), as a means of encouraging schools to encourage their students to pursue rigorous coursework.
- Other states tracked students one or two years beyond high school to develop measures of “transition to adult life,” such as the percentages of graduates who were enrolled in two- or four-year colleges; who were engaged in full-time work, part-time work, or military service; or who were unemployed.

➤ **Many states included additional subjects as part of their accountability frameworks.**

- In almost all cases, the additional subjects were those often called “core courses,” such as social studies or history in addition to math, reading, and science. In rare cases, additional accountable subjects include civics, economics, or geography.
-

➤ **Growth models were used in 15 of the 20 states surveyed.**

- Those 15 states calculated the growth in individual students' test scores from one year to the next. These student-level scores were then aggregated to the student subgroup and school level for school accountability ratings.
- Ten of 20 states calculated the average growth in groups of students' test scores from one year to the next by comparing, for example, fourth grade proficiency rates in 2010 to fourth grade proficiency rates in 2009 within the same school.
- In most states, growth alone does not determine a school's score; school ratings are typically determined by both proficiency rates taken from a single point in time that are compared to an absolute standard (as required by NCLB) and growth over time.

➤ **Most states established uniform expectations for test score growth rates, but some also constructed *relative* standards for growth rates by comparing a student/school to "peer" students/schools.**

- Such a measure factors in the reality that students grow differently over time and that each school has its own unique demographic distinctiveness. In addition, some schools have more funding than other schools, and the quality of teachers can widely vary. Low-income schools are more likely to have inexperienced teachers who are new to the field.

➤ **Many states with their own accountability systems weighted student performance along a performance spectrum, rather than applying a single proficiency threshold.**

- In 11 of 20 instances, states created a performance index score that assigned increasing points to student scores by performance level. In most of these 11 states, points were awarded within four or five performance levels (e.g., below basic, basic, proficient, advanced) as an incentive for schools to move children up the entire test score distribution.

➤ **States used climate and other factors as part of their indicator systems, but not as part of the traditional accountability system.**

States used school climate and other measures associated with student success to help gain an understanding of many of the factors that are not evaluated by a standardized test, but have a direct impact in creating the environment students need for successful learning.

Those factors included:

- **Positive school culture**, including student and teacher satisfaction, academic challenge, engagement, safety, and orderliness.
- **Positive behavioral, emotional, and physical health outcomes**, including indicators of suspensions, expulsion, and physical health.
- **Information about discrepancies in resources** available to schools and students.

-
- **State accountability systems often provide information about the context in which schools operate, but those measures were rarely used for positive or negative consequences.**
- These measures range from school principal turnover to parent-teacher conference participation and school-level per pupil spending.
 - Michigan was alone among states in factoring these measures into rankings. One-third of the basis for a school's letter grade was derived from a rating by school teams on a 40-item set of indicators that included both school process and performance measures.