accountability

building a Workable ACCOUNTABILITY SYSTEM

Hold schools more accountable for results and teachers and administrators will do everything in their power to ratchet up student learning. That theory underlies the recent proliferation of statewide accountability systems through which policymakers hope to ensure that education reform takes hold and yields the desired outcome of high achievement for all students. Coming up with an effective system is not an easy task. But much has been learned from experiences of such "reforming" states as California, Kentucky, and Texas — knowledge that can help guide policymakers as they consider their own state's accountability practices. This brief identifies seven key decision points on the road to a workable accountability system.

What are the primary goals?

Loading on too many goals undermines chances for success. Given pressure to be comprehensive, one option is a phased approach: stage one targeting the highest priority goal (e.g., improved reading/writing literacy) and other goals (e.g., improved mathematics literacy) introduced in subsequent stages. Priorities should be regularly reassessed.

What indicators should be included?

Fair and effective systems require multiple measures. In the category of *assessment indicators*, many states use both criterion- and norm-referenced tests to gauge how students are progressing relative both to state standards *and* to students nationwide, respectively. *Non-assessment indicators* include attendance, graduation, and dropout rates, as well as such measures as percentage of graduates enrolled in post-secondary education. The use of multiple measures requires careful planning: Indicators must be available (e.g., a standardized test aligned with state standards), valid, and reliable.

Which students should be tested, and when?

Given the need for a full picture of school performance, care must be taken in deciding who is tested, in what subjects, and when. The systematic exclusion of any student population could jeopardize the students' academic achievement by allowing schools to focus their efforts on those who "count" for accountability purposes.

- *All grades versus selected grades.* Considering the issues of testing burden and efficiency, assessing every student in each academic content area at every grade level can be excessive. States might better elect to assess different content areas at different grade levels, using a combination of norm- and criterion-referenced tests (e.g., test English language arts in grades 3, 6, and 9, and mathematics in grades 4, 7, and 10).

- *New or mobile students.* Title I and legal provisions in some states allow the exclusion of test scores from students who have been attending a particular school for less than a year. The rationale is that schools should not be held responsible for another school's teaching. A possible downside is the ongoing exclusion of highly mobile students (e.g., children of migrant farm workers).

- *Special populations.* Federal law requires testing of students with disabilities and of English Language Learners. Although decisions about how to accommodate special education students (e.g., Braille format, untimed test) are left to their Individualized Education Program committees, the

state determines whether results from heavily accommodated tests are comparable enough to warrant inclusion in the accountability system. Decisions about when to include English Language Learners are, increasingly, made at the state level, based on how long a student has been in the United States and served by an English as Second Language or bilingual program, and the student's score on an English proficiency test.

What accountability model best serves the purpose?

The most common model, found in Texas and New York, for example, relies on absolute benchmarks, comparing student performance data to state performance standards that are raised over time. In contrast, the model used by Kentucky and California examines performance over time, with the state setting target growth rates for each school based on demographics and starting points. Those farthest behind must generally demonstrate the largest annual growth. Kentucky's model sets growth rates for individual schools, but has also set a common 20-year goal for all schools — a compromise between judging against an absolute standard and encouraging schools to demonstrate continued growth.

What consequences can the system support?

The use of rewards and sanctions, key components in a number of systems, raises two important questions:

• Can the focused work required for continued improvement be sustained once the most easily effected changes

have been achieved? The risk is that instead of leading to improved learning for all students, consequences might, for example, move schools to target improvement efforts at a small percentage of students whose minimally improved test scores could boost a school's overall performance level.

• Are consequences based on reliable and valid indicators? If states cannot ensure that measured changes reflect accurate and meaningful differences in accomplishment or growth, the credibility of the entire system can be questioned. Generally speaking, including more indicators and data points increases reliability.

How can intended and unintended effects be evaluated?

A system may yield intended consequences (e.g., increased student learning) and unintended consequences (e.g., increased dropout rates). To maximize the former and minimize the latter, system adjustments may be needed. Continuous evaluation around five key questions is critical:

• Are the long- and short-term goals worthwhile, realistic, achievable?

• To what degree does the system support good instruction and student access to education; minimize corruption; affect teacher quality; and produce unanticipated outcomes?

• What costs are incurred and what is the necessary trade-off between quality and cost?

• What support do teachers and administrators need to implement the system (e.g., use assessment results to guide practice)?

• How will parents and the general public be informed about the system's goals and limitations?

What will be done to fix problems highlighted by the system?

States have an obligation to make sure schools have what's needed for success. Many have identified a cadre of support providers (e.g., distinguished educators, external evaluators), who help schools plan and implement needed change. Some have designated "improvement" funds for schools most in need of support.

Conclusion

How a state answers these questions affects the development, acceptance, and effectiveness of its system. All discussion must be infused with the awareness that an accountability system is a means to an end. Its goal is not to reward or punish, but to improve student learning.