

Work at policy level to broaden assessment

Summary

This strategy focuses on persuading the system leadership that the current narrow tests are a false economy, holding back progress by narrowing teachers' focus in a way that undermines learning. If the state standards reflect the national or international standards in any substantial way, the tests are also incompatible with the requirements of the NCLB Act, which requires state tests to be aligned with state standards.

Challenges addressed

State/district tests are too narrow

Test prep takes too much time

Key features

- Offer the system other assessment tools so that the total high-stakes assessment package is balanced across the system's goals, as represented by its standards.
- Use a recognized alignment process to show the current disparity between the current assessment and the standards.
- Point out the power of balanced high-stakes assessment in forwarding the school improvement process.
- Point out the modest cost in system terms of this change.

Background

US K-12 assessment's traditional reliance on multiple-choice tests is unusual, unique among advanced countries (e.g., those in the Organization for Economic Cooperation and Development, OECD). It arose historically through a variety of social pressures, and the dominance of a narrow behaviorist view of learning. This measurement perspective relied heavily on correlation to justify assessing only a small subset of the important kinds of performance. It ignored the effect of high-stakes assessment on the pattern of learning in classrooms.

Implementing the strategy

The core of this challenging strategy is to persuade the decision makers in your system that broader, more balanced high-stakes assessment is a cost-effective way to advance system goals for mathematical performance.

System and professional leadership will be only too aware of the challenge of persuading the decision makers to adopt any innovation, particularly if it involves changes in established processes and involves the reallocation of system resources. It needs to be developed and presented as part of a multi-year strategy for improved learning, and student performance on better tests.

The chances of success will be increased by assembling the usual elements of any such initiative:

- **Argument.** The essential elements are set out in this paper.
- **Evidence** to support the argument is important – though emotional responses may dominate, and must be taken care of first. There is strong evidence that, as every teacher knows, the kinds of items/tasks in high-stakes tests strongly influence

classroom learning activities. Types of performance that are not included in these tests are more-or-less totally neglected in many classrooms, which thus provide an education in mathematics that does not meet international, national (and in most cases state) standards. Systems such as *MAC*, that have built improvement around balanced assessment find no loss of traditional skills and substantial improvements in other important aspects of mathematical performance. Other references to research evidence are provided below.

- **Allies.** Prospects of success are always enhanced by ensuring that at least some members of key groups are in favor. Work to ensure that the superintendent, other administrators, colleagues in the professional leadership, principals, parents and, of course, teachers have positive attitude will, as always, pay-off. This will not succeed with every one, or every group. In this case those in charge of assessment in the system may well see it as both a criticism of their current practice, and a potential infringement on their "turf". Experience suggests that building a close collaboration between curriculum and assessment leadership is a great help, perhaps even essential, in achieving success. Teachers, whose professional lives will be most affected, must be assured of the value of the change, and that support will be available that will enable them to meet the new challenge.

As always with innovation, it may pay to start with a pilot project in a few schools. This will give experience with the various challenges of implementation, and build team capacity to support others. Only system-wide implementation as part of the accountability system will yield the full benefit of this strategy. It needs to be planned from the beginning.

How will this fit in?

This strategy will complement other aspects of your improvement program, and vice versa. *Professional development for and through performance assessment* proves highly motivating and enjoyable for teachers, even if the amount of time allocated for it is modest. *Balanced Assessment Classroom Packages* provide supplementary teaching and learning materials that will help teachers prepare their students for the broader range of performance that the tests, and standards, require. These will complement and enrich the main curriculum in the schools, whatever that may be.

Broader assessment will go on to provide reasons for expanded professional development, and adopting improved curriculum materials. (see *Assessment-led improvement*.)

Strengths

- **Improves teaching.** High-stakes assessment has a strong influence on what happens in the classroom – for worse or for better. Assessment that reflects the system's goals in a balanced way will forward its whole improvement program; narrow tests impede it.
- **Exemplifies the performance goals.** Good assessment tasks, released with scoring schemes and examples of student work, provide vivid exemplification of the performance goals of the system. They communicate more clearly to teachers and students than standards or curriculum, complementing these.
- **Cost-effective professional development.** The professional development support needed is relatively modest. A lot can be achieved in a few sessions, particularly when part of a continuing program. The scoring of such assessment by teachers is itself powerful professional development.

Likely challenges

- **Departs from tradition.** Some administrators see tests only in accountability and cost terms, understanding and thus caring little about what is assessed, or its effects on what students are taught. They are accustomed to, and thus happy with, short items and traditional multiple-choice tests that are simple and cheap to administer and provide widely accepted scores.
- **Seems to cost more.** The broad variety of longer tasks that are needed for balanced assessment have constructed responses and thus need human scoring, which is more expensive than automatic scoring. Administrators tend to think of costs in relation to the established assessment budget; in relation to total system expenditure, the contribution to system improvement of introducing broad and balanced assessment is highly cost-effective.

Overcoming these requires a process of educating those who make the decisions.

Tools

Standards-based assessment tools

Assessment that covers the full range of learning and performance goals of the system and its standards is essential for this strategy. Available assessment tools include:

Balanced Assessment in Mathematics: the tests. These tests, currently for Grades 3-10, are aligned with *Principles and Standards for School Mathematics (PSSM)*. They assess the broad range of performance that national and international standards require, with some associated classroom materials. The annual 40-minute tests at each grade are built from 5 to 10 minute tasks, which sample the broad domain of mathematical performance that the standards specify. They demand substantial chains of reasoning and non-routine problem solving. Materials are provided to support reliable local scoring – a contribution to teacher professional development. A scoring and reporting service is also available, as are practice tests. The tests are published by CTB/McGraw-Hill.

New Standards Mathematics Reference Examination. These tests, for Grades 4, 8 and 10, assess the broad range of performance that national and international standards require. They report scores for mathematics, with subscores for skills, concepts and problems solving. The tests are published by Harcourt.

Alignment analysis tools

While the mismatch between test items and the system goals for student performance is usually obvious, a formal analysis may be important in convincing decision-makers. Alignment protocols are available that support such analysis. The good ones go beyond lists of mathematical topics included to include other important aspects of performance, such as “depth of knowledge”, which short items rarely probe. It is important to use a protocol from a source independent of commercial test providers. Two that have high reputations are the *ACHIEVE Alignment Protocol* and the *Wisconsin Alignment Protocol*, designed by Norman Webb. The use of these protocols is a skilled matter. Training is available. The use of specialist outside consultants should also be considered – it also provides independence and credibility.