



Reliably Predicting Student Performance

A Guide to Choosing Assessment Solutions

About This Guide

We developed this guide with three goals in mind.

First, to help you better understand the widely varying types of assessment tools and methods that are available today, including the new approach of research-based predictive assessment.

Second, to explain why these tools should be combined to deliver the most accurate measure of student proficiency, as well as to effectively monitor student progress toward year-long learning goals.

And, third, to enable you to better question suppliers as you evaluate their products, so that you can put together the optimum assessment solution for your specific needs.

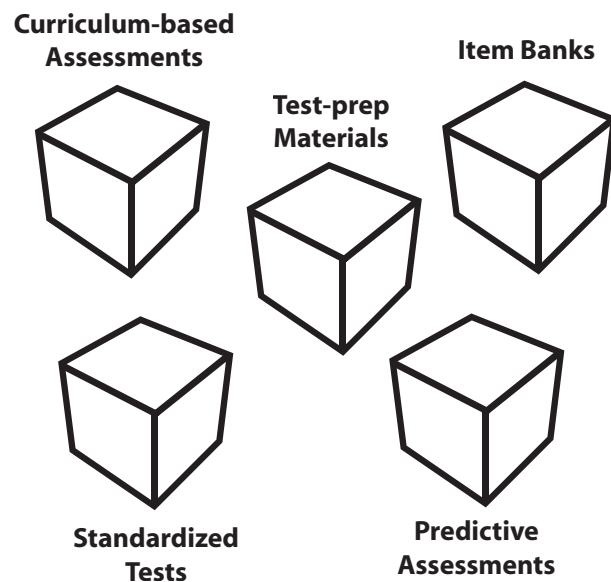
The Assessment Challenge

In today's world of state-specific standards and high-stakes tests, school administrators face a significant challenge. Once-a-year testing is no longer sufficient. It is now essential to track student proficiency throughout the year in order to help teachers focus their instruction and students to improve their performance.

Therefore, administrators are searching for better data tools and improved methods, such as formative assessments, to provide ongoing diagnostic information. Rushing to meet their need is a host of industry suppliers—literally dozens of vendors offering a variety of assessment solutions. As you see below, these solutions fall into five general categories. Which should you choose?

Making the Right Choices

The decision is less puzzling than it appears, because it is not a question of choosing a single tool. A combination is required in which the pieces fit—and work optimally—together. The next pages offer an overview of these tools and methods, and a glossary is provided at the back of this guide.



Assessment Tools and Methods

Here is a summary of the five most widely-used assessment tools and methods. Each has its advantages. Each helps complete the whole picture. This reinforces the need to put together the right combination of tools and methods to develop the most effective instructional strategy. At the same time, we believe that one tool is essential to any combination: the new approach called predictive assessment.

Test-prep Materials



- (+) Generally inexpensive
- (+) Students practice 'bubbling'
- (+) Focuses on test-taking strategies
- (-) No feedback during the year for students, educators or parents
- (-) Right or wrong answers only; no information to focus instruction
- (-) Items associated with a single subskill
- (-) Little or no information on skills or how they are measured
- (-) Little or no research basis; generally low quality

Item Banks



- (+) Flexibility to build tests for all purposes
- (+) Focuses on specific skills
- (-) Locally mapped rather than matched to state-specific standards
- (-) Success with teacher-created tests does not predict NCLB proficiency
- (-) Uncertain research to establish validity and reliability
- (-) Test items are not statistically designed to the difficulty level of the state test
- (-) No common information to share in grade-level teams to help redesign instruction

Curriculum-based Assessments



- (+) Feedback on each subskill tested
- (+) Instruction focused on making progress, paced by student performance
- (+) Subskills mastered in a building block approach
- (+) Provides useful instructional content

- (-) May or may not provide information to share with grade-level teams
- (-) Measures performance on subskills, but not on state-specific proficiencies
- (-) Doing well does not necessarily mean proficient
- (-) Provides a short-term view of learning goals, not a year-long view

Standardized Tests



- (+) Comparable, but different than the state-specific test from same supplier
- (+) Has similar reliability, validity and reports as the actual state test
- (+) Provides a benchmark of student performance

- (-) Usually the most expensive solution
- (-) Provides a progress point from only one test
- (-) Teachers don't get to keep and reuse the items

Predictive Assessments



- (+) Predicts mastery, proficiency and state-specific AYP
- (+) Teachers and students get quarterly feedback on annual learning goals
- (+) Matches diagnostics to each state's high-stakes test
- (+) Teachers can see NCLB results before they administer the state test

- (-) Subskill information is available, but measured as part of the state standard
- (-) For students who are two years behind, additional diagnostic information is needed

Assessing Assessment: Evaluating Your Needs

The following questions and scoring guide will help you evaluate how well the assessments you use now are meeting your goals for student achievement.

1 How well are most of your students achieving on the state high-stakes test?

- A. Low B. Average C. High

2 In the average class in your district or school, which describes your students?

- A. While some are high performing, many struggle to learn and meet proficiency.
- B. Most classes have a mix of low, average and high performing students.
- C. While some are low performing, most perform well and our teachers work to continue to challenge their students.

3 Which scenario best describes the faculty in your district or school regarding the use of assessment data?

- A. Many say they don't have time.
- B. We have a mix. Some are experienced. They follow state standards and use state and curriculum-based data. Others think we test too much and that data analysis keeps them from teaching.
- C. Most are seasoned educators who use a variety of data and analysis tools to inform instructional placement and strategies.

4 Are you using data to inform instruction?

- A. We know we need to start using data, but it is a challenge for us to interpret.
- B. We are beginning to use data. Some teachers interpret data easily, others are new and/or don't see its value in revising instruction or identifying struggling learners.
- C. The leadership and faculty are experienced in using data-driven approaches to systematically inform instructional placement and strategies.

5 What is your goal for teachers in using assessment data?

- A. We want to pre-test students and measure progress throughout the year. We want to drive our instruction with data in every class.
- B. Our teachers know the curriculum well and develop good tests, but want an outside measure to be sure the students are on track.
- C. We just need students to practice immediately before the high-stakes test.

6 What is your goal for parents in using assessment data?

- A. We believe that providing assessment data to parents will give them reason to communicate with us more often and more meaningfully.
- B. We want parents to better understand what their child needs to learn, using easy to interpret data.
- C. Our parents aren't interested in looking at data.

7 What is your goal for students in using assessment data?

- A. We want our students more engaged in those curricular areas where they need improvement.
- B. Give students feedback and positive reinforcement when they reach individual target goals.
- C. Our students won't understand how data could inform their learning activities.

8 What data requirements do you have as part of your School Improvement plan(s)?

- A. Our staff meets and develops measurable goals based on identified learning gaps using prior state/district tests and/or curriculum based tests.
- B. We use data for special education IEPs, but teachers make their own plan.
- C. We don't have data requirements in our school improvement plan.

9 What help do your teachers need to better understand and use the measures generated by standardized tests?

- A. They need guidance to review our in-house data and the state's high-stakes test.
- B. They need help finding test items that match our state standards and are at the same level of difficulty as the state's test.
- C. They don't need much help. We effectively use a variety of state, local and vendor product reports to drive all of our school and teacher plans and have seen improvement.

10 Are high-stakes test scores an issue at your school?

- A. Yes, predicting where each student needs to focus, getting our teachers to build on each others strengths and knowing what special program resources to use is a challenge.
- B. Each year, the issue is a bit different – overall, we're making progress in monitoring each student and provide them with appropriate instruction.
- C. It isn't a big issue for us - our teachers get 90% of our students to expected performance levels and we are working on the rest!

What Are Your Needs?

3 points for each A
2 points for each B
1 point for each C

Total points shown right.

1-9: You are satisfied with your current approach to assessment. You may want to consider using some web-based free 'released items' to support your teachers.

10-19: You are using assessment data with some effectiveness. Compare your current assessment approach to research-based predictive assessment to see if that method can help you improve your progress toward year-long learning goals.

20-30: You fit a profile of schools and districts that have successfully improved student scores on high-stakes tests through improved use of data to drive instruction.



Predictive Assessment

We stated earlier our belief that predictive assessment is an essential component of the most effective overall solutions. To help you evaluate products from different suppliers, here are the five key requirements an assessment must satisfy in order to be truly predictive.

1 Match Specific State Standards

The questions must match every skill or standard for the state's curriculum, and document the state-specific subskills being addressed.

2 Match the Difficulty-Level of the State High-Stakes Test

The questions in a predictive assessment are at the same level of difficulty as the test the state will administer to its students. Importantly, the difficulty level must match for every skill and every subskill.

3 Match the Range of Questions in the State High-Stakes Test

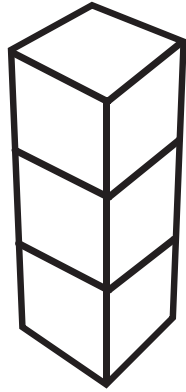
A predictive assessment includes a representative sample for every skill measured in the state's high-stakes test, not just some of them or skills being taught in a given week.

4 Predict Each Student's Proficiency According to State-Specific Definitions

A predictive assessment forecasts individual student performance based on state-specific definitions of student proficiency. Results from actual test scores are compared to the forecast to refine the assessment. Administrators can be confident in the forecast.

5 Improve Student Achievement

A predictive assessment generates data from large student populations. The data drives curricular strategies that improve student learning and performance on state tests. Improvement is verified through research.



Putting It All Together

The good news for administrators is that there are tools and methods available to help them meet the assessment challenge. And, while no single approach is a complete solution, they can work together in combination to ensure the best possible assessment solution for your specific needs.

To help you put the right combination together, we close by listing the six key criteria that every assessment product should satisfy. Use these criteria when evaluating a product, and make sure that the supplier can demonstrate that their offering is:

- ✓ **Affordable**
Requires minimal staff time, training, equipment, and paper.
- ✓ **Easy**
Quickly implemented with ready access to ongoing support.
- ✓ **Understandable**
Shows results clearly, without the need for statistical analysis.
- ✓ **Flexible**
Works well online or on paper.
- ✓ **Compatible**
Data integrates easily with your other systems.
- ✓ **Research-based**
Validity and reliability statistics with scientific evidence of results.

Glossary

Summative Assessments - Measures or “summarizes” a student’s knowledge at the end of a learning experience providing information about final student performance. The intent is to assign a score or record a grade.

Formative Assessments - Monitors student progress towards goals within a content area. The intent is to guide the instructional and learning process based on the feedback teachers and student receive from the assessments.

Norm-Reference Tests - Student performance is compared against the performance of a group of students, called the “norming group”; that is intended to represent the target population. Scores are generally reported using National Percentiles that rank how a student performed relative to other students.

Criterion-Reference Tests - Student performance is compared against a common standard or expected level of performance. The test is designed to identify if the student has met or failed to meet a predefined body of knowledge.

Diagnostic Assessments - Intended to serve as a basis for remedial instruction, these assessments measure specific skills or abilities within a broad area of knowledge.

Test-prep Materials - Black-line masters and released or sample items that give students practice with the test environment and item format. This solution is intended to provide test-taking strategies for students.

Item Bank Assessments - A pool of test questions aligned to a set of standards. Usually delivered online, the solution allows the user to select specific targeted standards to assess student performance on the chosen skill. The same pool of items is frequently used across different states and remapped, often without a human review, to each state’s standards.

Curriculum Assessments - Assessments are related to the product’s own content and curriculum detailing a specific scope and sequence of skills. The purpose is to guide the delivery of the program’s instructional content and evaluate student mastery of individual skills.

Standardized Tests - Produced by major testing companies with a sufficient national sample to offer a nationally normed test allowing comparison of students from one state to another. Typically these tests are not aligned to state standards unless the company has a publicly bid contract to provide its test. These tests provide a benchmark of student performance at a single point in time.

Predictive Assessments - A research-based system designed to forecast a student’s performance on another test measuring similar skills and abilities. Using reliable and valid measurements, predictive assessment utilizes a formative approach to monitoring and improving student performance on summative measures such as year-end or end-of-course evaluations.

Adequate Yearly Progress (AYP) - A measure of year-to-year student achievement of statewide high stakes tests mandated by NCLB. Each state sets the minimum levels of improvement that districts and schools, as a whole and as subgroups, must achieve within specific time frames in order to meet 100% proficiency by 2013-2014.

Scale Scores - Scores are reported using a single scale with intervals of equivalent size allowing for comparisons across different groups of student regardless of time of year or differing group characteristics. Such comparisons across time or between groups is not possible using Grade Equivalents or National Percentiles.

Grade Equivalents - Identifies student achievement according to the performance of the typical student of the norm group who had completed that grade level and month of the total school year. It does not indicate that a student has mastered all skills through that grade level.

National Percentiles - The score indicates the percent of students in the norm group that scored below the student or group that was assessed. The intent is to compare the performance of a student or group to how other students or groups performed nationally. Percentiles cannot be averaged or aggregated.

Normal Curve Equivalent (NCE) - Equal interval scales that are roughly equivalent to percentiles with a range of 1-99 and a mean of 50. Because they are composed of equal sized units, the scores can be aggregated and averaged. NCEs identify how many students in the norm group out of a hundred scored below the assessed group. The intent is to measure gains or losses of a particular group and not an individual student.

About the Authors

Michael K. Smith, Ph.D., is a well-known author and psychometric researcher who serves on the adjunct faculty at the University of Tennessee. Smith has twenty years of experience developing and testing assessments for teacher certification. His latest book, scheduled for release in 2005, traces the history and impact of standardized testing. Smith also serves as a consultant to ThinkLink Learning, a Nashville, TN based provider of assessment solutions for K-12 schools.

William Dycus, EdD., is psychologist and educator who specializes in strategies for helping low-performing schools meet standards for student achievement. Dr. Dycus has been a teacher, school principal, and administrator, and is currently the Coordinator of Psychology, Metropolitan-Nashville TN Public Schools, where he manages a staff of 73 to provide special education services to 70,000 students. He has served on the Board of Directors for the Association for Supervision and Curriculum Development (ASCD) and worked in many other professional and leadership capacities, writing and presenting to education colleagues.

Jacqueline B. Shrago, is a former educator who has for twenty years developed and advocated technology to support education. She developed the concept of a formative assessment using a summative approach and launched ThinkLink Learning to build research-based technology tools that are easy for teachers to use and understand. Shrago has been an advisor to the National Governor's Assn., Council of Chief State School Officers, and CEO Forum. She has received recognition for her work on the use of the Internet in education from the Nat'l Information Infrastructure and Smithsonian Computer World Awards.

Sources

Black, P., Harrison, C., Lee, C., Marshall, B., & William, D. (2003). *Assessment for learning: Putting it into practice*. Maidenhead, England: Open University Press.

Foos, P.W. & Fisher, R. P. (1988). Using tests as learning opportunities. *Journal of Educational Psychology*, 80(2), 179-183.

Fuchs, L. & Fuchs, D. (1986). Effects of systematic formative evaluation: A meta-analysis. *Exceptional Children*, 53(3), 199-208.

Jones, M. G., Jones, B. D., & Hargrove, T. Y. (2003). *The unintended consequences of high-stakes testing*. Oxford: Rowman & Littlefield.

Lane, S. (2004). *Validity of high-stakes assessment: Are students engaged in complex thinking?* Paper presented at the meeting of the National Council on Measurement in Education, San Diego, CA.

Popham, W. J. (1999). *The truth about testing*. Alexandria, VA: ASCD.

Pellegrino, J. W., Chudowsky, N., & Glaser, R. (Eds.). (2001). *Knowing what students know: The science and design of educational assessment*. Washington, D. C.: National Academy Press.

Sacks, P. (1999). *Standardized minds*. Cambridge, MA: Perseus.

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, 18, 119-144.

Safer, N. and Fleischman, S. (2005) *How Student Progress Monitoring Improves Instruction*, Association for Supervision and Curriculum Development, 81.

Smith, M. and Shrago, J. (2005). *Online Assessment in K-12 Classroom: A Formative Assessment Model for Improving Student Performance on Standardized Tests*. In Scott L.H. (Eds) *Online Assessment and Measurement, Volume II: Case Studies from Higher Education, K-12, and Corporate (Chapter XIII)*. IDEA Group Publishing.



This guide produced with support from:

ThinkLink Learning™

2409 21st Avenue, South, Suite 202

Nashville, TN 37212

(866) 814-6685

www.thinklinklearning.com

©2005 ThinkLink Learning, Inc. All rights reserved.

Printed in U.S.A.