

Accountability: Setting Expectations, Measuring Performance and Providing Support

Parents, the public, educators and policy makers want to know how schools, teachers and students are performing. "Accountability" is the term people commonly use to express this demand for results and desire to hold states, districts, schools and students responsible for academic performance. Accountability systems that press for improvements in student achievement have grown out of standards-based reforms as ways to ensure students are meeting standards and, where they aren't, make sure schools are helping them develop the needed knowledge and skills. (See January 2000 *Parent Press* on standards.)

States and districts use assessments to collect the data on which to gauge improvement and base decisions about rewards or assistance. Assessments are ways of collecting data to know if students, schools and districts are making progress. Tests are commonly used but student work, grades, attendance rates and observations are other measures. Accountability systems, then, define how the data will be judged and what indicators constitute progress.

There are several levels of accountability. The federal government generally holds states accountable for overall performance of the schools in that state. States hold school districts, and sometimes individual schools, responsible. Districts hold students, and sometimes teachers and principals, responsible for performance. Principals sometimes hold teachers responsible, and in classrooms, teachers hold individual students responsible.

Generally, the data from assessments are made public. This helps parents and communities hold schools responsible and may encourage schools and districts to make improvements. Based on the data, states make decisions about rewards for improvements, often in the form of cash and recognition. States may enforce consequences for inadequate performance over a period of time by replacing school or district leaders and sometimes a group of teachers, or by taking over the governance of a district or school. More states are finding ways, as a part of their accountability system, to provide targeted assistance to struggling schools before issuing these tough sanctions. This assistance often includes additional

money through school improvement grants or technical assistance from teams of people from the state department of education.

Evaluating progress

At each level of accountability, different kinds of assessments are used to measure performance. These different types of measurements can be used for various purposes by states, districts, schools and teachers.

Although tests are the most common—and often the highest profile—form of assessment, there are different kinds of tests. The differences generally lie in what kind of information they produce about performance.

When determining the type and purpose of a test, it's helpful to ask:

- Does the test give a wide-angle picture of how students are performing in comparison to their peers? OR
- Does it provide a close-up picture of a student's knowledge and skills in a certain content area?
- How is the test scored?
- How are the results reported?

Standardized tests, which are given to large numbers of students in uniform conditions, are used to measure and compare overall student achievement at district, state and national levels. There are two kinds of standardized tests: "norm-referenced" and "criterion-referenced" tests.

Norm-Referenced Tests

Traditionally, in this country, tests have been used as tools to select and sort students. In this sorting mindset, most states have relied primarily on norm-referenced tests—like the Iowa Test of Basic Skills, Terra Nova, or Stanford Achievement Test-9th Edition (SAT-9)—to hold schools and districts accountable.

These tests, which compare students to one another based on a "norm," are limited in what they can measure about what a student knows and is able to do.

Ruth Mitchell of the Education Trust, a Washington D.C.-based education advocacy group, explains that a test that is scored on a bell curve "insists that *no matter how good the teaching may*

be, 50% of the children will always be below average and 50% above."

Norm-referenced tests have become an "entrenched" part of educational accountability because they offer such clear comparisons, says Cris Gutierrez, an educator in Santa Monica, Calif. who works with several national education groups on teaching, accountability and peace education. But norm-referenced tests are limited in how they measure student achievement toward an agreed-upon set of standards.

Criterion-Referenced and Standards-Based Tests

Criterion-referenced tests measure individual students' performance against a pre-defined body of knowledge—rather than compare students to their peers. In the era of standards-based reform, the criteria measured are standards, and consequently these tests are now called stan-

dards-based tests. Many states are developing their own standards-based tests to measure students' performance toward the knowledge and skills defined in the state academic standards.

Results of standards-based tests generally report levels of how well the student has mastered the content: basic, satisfactory, proficient, advanced. From this kind of reporting, schools and teachers get data about the specific kinds of knowledge and skills students have mastered or may still need help learning. These data enable teachers and students to gauge progress toward meeting standards.

The January 2001 issue of *Quality Counts*, titled *A Better Balance: Standards, Tests, and The Tools To Succeed*, published by *Education Week*, reports that in 2000, 40 states were using standards-based tests in English at the elementary, middle and high

school levels, and 34 were using such tests at all three levels in mathematics. These tests are 'aligned' with state standards.

Examples of state standards-based tests include the Texas Assessment of Academic Skills (TAAS) and the Massachusetts Comprehensive Assessment System (MCAS).

The National Assessment of Educational Progress (NAEP), a criterion-referenced test, has been administered by the U.S. Department of Education regularly since 1969 to schools and students who volunteer to participate. Called "The Nation's Report Card," the NAEP measures student achievement nationally and by state (but not by student) in various academic areas by analyzing data from a sample of student scores.

But even when tests match the standards, they may still lack rigor or be too limited to measure the full range of what students know and are able to do, according to an independent review conducted by Achieve, Inc., a non-profit group formed by governors and business leaders to promote standards-based reforms, as reported in *Quality Counts 2001*. On the whole, Achieve found that state standards-based tests don't match the corresponding standards as closely as they should. Nor do these tests measure a wide range of content or as much demanding knowledge as they could.

Multiple measures of progress

In the field, explains Gutierrez, practitioners are trying to get a better handle on "multiple mea-

(continued on page 2)



PARENTS FOR PUBLIC SCHOOLS
NATIONAL OFFICE
1520 North State Street
Jackson, MS 39202-1645

NON-PROFIT ORG
US POSTAGE
PAID
Permit #603
Jackson MS

Letter from the Executive Director

April 2001

Among the various reform issues *Parent Press* has addressed in recent months, testing and accountability are perhaps the most controversial. This is true, in part, because testing has become a lightning rod for holding persistently low-performing schools accountable. Policy makers and schools are consumed by it. And yet, the testing debate need not be the only driver of accountability and reform. A disquieted public is also concerned about equity, improved teaching, and the continued support for high standards.

When asked about testing, most parents I talk with are eager to know how their students are performing, but lament that so much classroom time is spent on test preparation and test taking. We are confused about what's being measured and what the tests tell us about individual achievement. We wonder why test scores don't match class grades. We are concerned that, although tests are finally providing much-needed data, they are not necessarily bringing about improvements in teaching or learning. We want to know what other sources of information, if any, schools are using to determine how our students and teachers are performing. You'll find the answers to these questions in this issue of *Parent Press*.

Every effort has been made to present within these pages a comprehensive and balanced picture of these related and politically charged subjects. This issue is filled with resources that can help parents and educators tap into useful information about their own local landscape of testing and accountability. For parents to effectively support student learning, we need to understand the differences between tests, how tests are used, and how tests are (or should be) connected to curriculum and standards. And as momentum builds for improved accountability systems, we also need information about a whole set of conditions that support learning: high expectations, school readiness, professional development, and community involvement, to name a few.

I hope that what is included here will give parents new tools and insights to understand the role testing plays in a *balanced* accountability system—a system that uses testing as *one* measurement among others to guide school improvement efforts. As PPS chapters gain knowledge about these and other important reform



issues, they will become articulate and credible forces for school change. Testing and accountability are complex issues, but parents have a right and a responsibility to be a part of the debate.

Kelly Allin Butler,
PPS Executive Director

PARENTS for PUBLIC SCHOOLS
NATIONAL OFFICE
1520 North State Street
Jackson, MS 39202-1645
1-800-880-1222
www.parents4publicschools.com

Executive Editor: Kelly Allin Butler
Project Editor: Kris Kaiser Olson
Strategy, Research and Writing:
Collaborative Communications
Group
Design: Marie Owen,
Owen Advertising Studio

Accountability (continued from front page)

“sures” of progress. They are exploring, but have not yet perfected, ways to pull from two or more forms of assessment to make judgments about progress for students or schools. Waco Independent School District's promotion policy is an example of how a district uses several factors to gauge progress, diagnose needs for improvement and make decisions about consequences. (See story, page 3.)

Some states are using attendance rates and dropout/graduation rates in addition to test scores as the criteria for defining success or indicating need for assistance in schools. Rhode Island's School Accountability for Learning and Teaching (SALT) is an example of a state accountability system that uses a variety of ways to evaluate performance. In addition to test scores, the SALT program provides teams of educators who visit schools once every five years. The program also conducts surveys of teachers, other school employees and parents about what's happening in classrooms and in students' lives.

District- and school-level measures of student progress

Districts and states hold students responsible for performance with grades, courses taken, progress on their classroom work and attendance, in addition to test scores. These criteria—measured at the district or school level—can be used to make decisions about graduation, promotion from grade to grade or the need for extra academic support.

At the risk of over-testing students, many districts develop their own standards-based tests. These local tests are particularly useful in states that rely on national norm-referenced tests, or that don't provide highly useful, timely data from the state test.

For example, the Long Beach (Calif.) Unified School District has developed its own Open-Ended Mathematics test to measure students' mastery of math standards. This kind of test, which asks students to provide answers, as opposed to recognizing answers on a multiple-choice test, is called a “performance assessment.”

According to Dixie Dawson, mathematics curriculum leader in Long Beach, the open-ended tests, which are part of a mathematics portfolio in grades 1 through 8, “make it really clear what we want kids to know at each grade level. With this clarity, teachers make sure they give assignments that help students master the content called for in the standards and the assessments.”

At the school level, portfolios and student work can provide data that show progress and places for additional learning by students or adjustments in in-

struction by teachers. **Portfolios**, or collections of student work and results of assessments, provide rich information for students and teachers about what and how well a student is learning. Portfolios, though, are sometimes cumbersome in the way the components are defined and in the way they are implemented in the classroom.

When teachers participate in a structured, regular process of **looking at student work** and their own assignments, they use student work as data to determine whether they are holding high expectations for students and what corrections they need to make in their assignments or teaching methods.

Reporting progress

Quality Counts 2001 reports that 45 states publish annual report cards on schools and 34 publish reports on districts, as the most common public reporting tool for state accountability systems. Several states publish these report cards on their Web sites.

Jim Watts, in “Getting Results with Accountability,” a publication from the Southern Regional Education Board, writes that state report cards should focus on student achievement and educational results and be useful for school improvement, among other criteria.

“Good reports,” Watts writes, “focus on how students are performing and reflect standards and assessments that are part of the overall accountability system.” They should show trends and provide timely and accurate information.

When states report disaggregated test data—which means they break apart big data chunks by student and school characteristics such as ethnicity or participation in free/reduced lunch programs—educators and parents are better able to spot achievement gaps and areas for improvement among students and teachers, and in specific schools.

Watts warns, however, that when states and districts use disaggregated data, they should not let demographic factors “excuse students' poor performance or be used to lower expectations.”

Using data

Among the different kinds of tests, it's crucial to discern how a school, district or state *uses* test scores. When investigating how data are used, it's important to consider:

- Does the use of data match the purpose of the test?
- Are scores used to identify which schools aren't performing, and define the kinds of help that school's teachers and principal need to improve?
- Is one test used to make high-stakes decisions?

High-stakes decisions

Many states and districts have adopted a high-stakes approach

to testing. “High stakes” means that performance on one test is used to make an important decision for a student or school, such as whether a student is promoted to the next grade or graduates from high school, or a school is subject to harsh consequences. According to *Quality Counts 2001*, students are more likely to be subject to high-stakes decisions than individual teachers or principals.

High-stakes tests put special education students at a particular disadvantage, especially when they don't have appropriate accommodations in the test format or environment. Recently, in a settlement of a class action lawsuit, the Oregon State Board of Education agreed to modify the state's current testing system based on an expert panel's recommendations to ensure that students with learning disabilities are treated fairly on standardized tests.

Diagnosing areas for improvement

Ideally, data is used as a diagnostic tool for improving teaching and learning. State test data can provide information about needs for improvement when data is reported quickly and broken down, or disaggregated, according to classroom, subject areas and key groups of students. When used well, data collected from evaluation measures should drive changes in professional development and define the other kinds of supports and resources that schools and students need to improve their performance.

“The key to a good diagnostic measure,” says Ellen Pechman, a K-12 evaluator with the Washington D.C.-based Policy Studies Associates, “is that it is done at the level of student or classroom and can be scored on a quick turnaround, so that the data comes back quickly and in formats that teachers can use.”

Diagnostic assessments are more commonly, and easily, used at the classroom level. For instance, first grade teachers can conduct “running records” of individual students' reading skills four times during the school year to measure progress and know the areas in which a student needs extra help.

The future of accountability

Making comparisons

Policymakers and the public have gotten used to making comparisons in order to measure educational progress. The results from the Third International Mathematics and Science Study (TIMSS), conducted in 1995 and again in 1999, are frequently used to make judgments about the U.S. education system based on comparisons with other countries.

The 1999 TIMSS study focused on the math and science achievement of eighth-graders in 38

(continued on page 3)



PARENT
PRESS

Parents & Accountability

Greater accountability for schools allows parents to be more involved than ever before. Accountability generally involves a public reporting of results in the form of school report cards and test scores. But too often parents aren't sure what all of the information means and what to do with it.

Lynn Olson, a senior editor at *Education Week*, advises that parents understand *all* the performance expectations for their local schools. "Parents need to be clear about what schools, teachers and students are being held accountable for, whether it's having all students pass a test, closing the achievement gap, course-taking patterns, high school completion or graduation rates."

Parents also have a role to play in making sure accountability decisions aren't arbitrary. "Parents can press to make sure that if there are rewards and consequences for schools and students, the basis for deciding whether schools receive rewards or sanctions is clearly defined," Olson says.

For the public to be able to examine accountability systems, they need access to information about achievement levels and what tests the state or district are using to measure achievement. However, while school report cards can be key sources of information for parents and the public, parents generally aren't getting the information they need or want to understand their state's or district's accountability and testing systems.

Parents need better information

Survey data indicate that parent and local advocacy groups can play an important role in educating parents about the kinds of tests used, how they are used, and what that means for students and schools. A series of recent surveys indicate that, across the country, parents generally support testing, but express a lack of knowledge and concern about high-stakes testing.

In an August 2000 nationwide survey for The Business Roundtable, Belden Russonello & Stewart, a Washington D.C.-based polling firm, found that three-quarters of the American public supported the principle of requiring students to pass a reading and math test for promotion from 4th to 5th grade, and two-thirds agree with the policy of requiring students to pass a test of their knowledge and skills to graduate from high school.

Yet, parents think that tests should be balanced with grades and teachers' assessments to determine readiness for promotion or graduation. Nearly half of the parents participating in a June 2000 survey by the Association for Supervision and Curriculum Development found inconsistencies between their child's standardized test scores and their report cards.

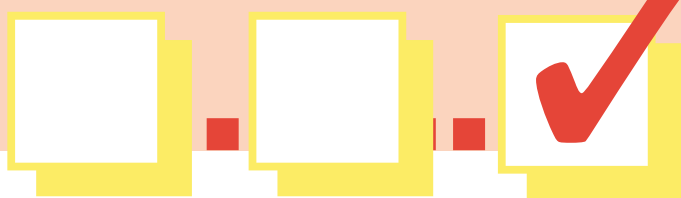
The survey, conducted by Harris Interactive, also reports that half of the parents surveyed were unsure of or do not know what the state standardized tests measure, nor do they feel equipped to help their children on standardized tests. Parents said they need more specifics about (1) content, (2) format, (3) grade levels of assessments, and (4) skills measured on the tests.

School report cards

Since school report cards, required in 45 states, are meant to help schools be accountable to the public, the public needs to be able to understand them. Parents can help push the district or state entity that publishes them to explain the data and concepts in parent-friendly ways, and provide a context in which to understand the data. Parents can also insist that school district data be reported in disaggregated formats that break up the data according to specific schools, groups of students, or even classrooms.

Parent or community groups can also prepare their own parent-friendly public reporting of data. The Pittsburgh Council for Public Education collects data from the school district and publishes it in a clear format with graphs and charts demonstrating key indicators of student and school performance.

Another group, the Public Education Fund in Providence, RI, involved parents in developing "school progress reports" by convening focus groups and using what they said they wanted to know about their schools to develop report card templates that schools could easily fill in and publish. (Examples of both of these report cards are available on the Web. See Resources on page 4.)



Accountability

(continued from page 2)

countries by testing students and analyzing other factors that affect learning. The 1999 results show that U.S. students performed near the middle of the international average in mathematics, with achievement in science relatively better. In addition, TIMSS shows that U.S. students' performance levels decline as they move from elementary school to high school.

Traditional norm-referenced tests have also provided the kinds of comparisons nationally and within states that policy makers are looking for. So, as the use of criterion-referenced tests grows, there is some concern that the public will lose the ability to compare achievement levels. Jennifer Vranek, the director of

Achieve's efforts to work with states to evaluate their standards and tests, contends that standards-based tests can also provide comparative data. "If a student is 65% of the way to meeting state standards, that can be compared to an average of 42% of the kids in the same state that are on the way to proficiency." This kind of score would indicate that the student is performing better than many of his peers across the state.

Heightened role of testing

The President's education proposal increases the stakes for testing by requiring states to develop tests that measure "basic skills and essential knowledge" in reading and math and to administer those tests every year to students in grades 3 through 8. Regardless of the final education

legislation passed by Congress, elected officials and educators will continue to seek ways to measure academic performance to obtain information they need to make judgments about performance and assign rewards, consequences and assistance.

If the President's proposal for annual state testing survives Congress, Vranek suggests that parents advocate in three areas related to these new testing systems: "Parents should be advocating that (1) the tests are standards-based; (2) parents and schools can get results back in ways that help them; and (3) states aren't sending mixed signals by combining the use of 'off-the-shelf' norm-referenced tests in some grades and the state standards-based tests in other grades."

A new model of accountability

In a model of accountability developed by Cross City Campaign for Urban School Reform, accountability needs to be "reciprocal," which means that everyone, including parents and community, accepts—and shares—responsibility for student achievement. According to Cross City's model, "parents, students, and community members are recognized as essential partners and accorded full respect." They should "help design, implement, and evaluate improvement plans."

In a similar effort, a consortium of national groups, including the National Center for Research on Evaluation, Standards, and Student Testing (CRESST), the Pew

Forum for Standards-Based Reform at Harvard University and Achieve, are developing a model of "collective accountability." For example, explains Gutierrez, when parents call for and get better data showing students' and schools' performance, teachers should also benefit because that data will be made available to them as well. The model is based on the belief that a wide range of groups and people looking for quality data and high levels of performance creates a "collective call" for improvement.

In these new approaches to accountability, there are places for parents and communities to engage in efforts to hold states, districts, schools, educators and students accountable.

GLOSSARY

Accountability is the notion that state education systems, districts, schools and students need to be held responsible, or accountable, for their performance. Accountability policies define how schools or students are judged and what consequences or rewards they will be subject to based on how they perform.

Alignment in a standards-based system is used to indicate that a test, curriculum or another aspect of education, is aligned to or matches the standards. When a test measures the skills and knowledge outlined in the state standards, it is aligned to the standards.

Assessment is the measurement of a student's skills or knowledge in a subject area. Tests and student portfolios are both forms of assessment. Performance-based, or authentic, assessments require students to apply their knowledge in hands-on tasks, such as writing an essay or conducting a science experiment.

Criterion-referenced tests measure a student's performance based on a set body of content. State tests that match state standards (standards-based tests) are criterion-referenced tests.

Data, gathered from student assessment or other factors in schools, provide information about student achievement levels. When data are reported, or disaggregated, according to demographic characteristics and reported in a timely manner, they are helpful in identifying areas for improvement among students and teachers.

High stakes tests are used as the only criterion to make an important decision for a student or school. The test score may be used to decide whether a student graduates or a school is subject to harsh consequences, such as replacing the leadership and much of the staff.

National Assessment of Educational Progress (NAEP), sometimes called "the nation's report card" is a test administered by the U.S. Department of Education. The test is voluntary and does not report scores for individual students.

Norm-referenced tests compare students' performance to a "norm" group of students. Scores are reported on a bell curve, where 50% of the students who take the test fall below average and 50% fall above.

Performance assessments ask students to produce their own work in open-ended, short answer questions; with written essays; or by performing a math equation.

Portfolios are systematic collections of student work, self-reflection by students, and results of assessments throughout a course or class year.

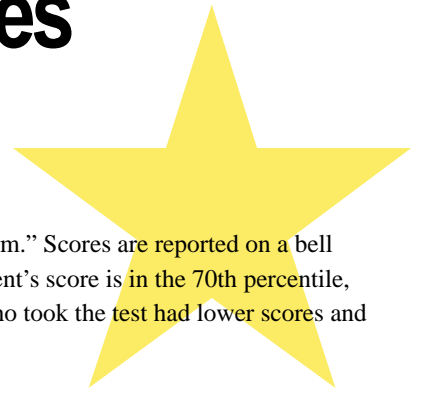
Reconstitution, as a "get tough" consequence for poor performance over time, refers to a district or state's action to replace a superintendent, principals or even groups of teachers in efforts to change the culture of a school or district.

Standards: *Academic standards* are clear statements about what students should know and be able to do in certain subject areas and at certain stages in their education. *Performance standards* describe how good a student's work must be, or what level of mastery is required, to meet the academic standards.

Standards-based reform is an effort to increase student achievement by setting high academic standards and then changing the way schools and classrooms work so that all students can achieve to those high standards.

Standardized tests are constructed so that the questions on the test and the way it is administered and scored are uniform, or standardized, for the large numbers of students taking the test. "Standardized" does not mean that the tests are aligned with standards.

Navigating the world of tests: What are the differences between them?



Norm-Referenced Tests

How they are scored: Compare students to one another, based on a "norm." Scores are reported on a bell curve using "quartiles" or "percentiles." If a student's score is in the 70th percentile, for example, it means that 70% of the students who took the test had lower scores and 30% had higher scores.

Examples: Iowa Test of Basic Skills, Terra Nova, Stanford Achievement Test-9th Edition (SAT-9)

Use: Developed by national test and textbook companies. They offer achievement data that compare students according to a norm selected by the school district. These tests do not tell you how much a student knows; they do report how much better or worse a student's performance is when compared to others judged on the same norm. For example, if school districts use an "inner-city" norm, the students are compared only to other students in inner cities, not to the population of all students.

Criterion-Referenced Tests

Criterion-referenced tests are now known as standards-based tests

How they are scored: Performance is reported according to the level at which a student achieves the standard. For example, a continuum of scores may be: basic, satisfactory, proficient, advanced.

Examples: Texas Assessment of Academic Skills (TAAS), Massachusetts Comprehensive Assessment System (MCAS), National Assessment of Educational Progress (NAEP), Third International Mathematics and Science Study (TIMSS)

Use: Measure how much a student knows about the standards, which are the criteria being tested, and measure students' progress toward attaining the standards.

Standardized Tests

"Standardized tests" are tests given in uniform conditions: the same questions are given, all students are allowed the same amount of time to do the test and the test is given in the same part of the academic year for all student who take it. "Standardized" does **not** mean that the tests are aligned with standards.

Types of Test Questions

Typically standardized tests have been made up of mostly *multiple-choice questions*.

Tests that measure progress towards academic standards usually have some, although not all, *performance assessment questions*. These questions ask students to produce their own work in open-ended, short answer questions; with written essays; or by performing a math equation. A performance assessment question for an 8th grader might ask:

Nuclear energy can be developed for peaceful purposes. What is one argument in favor of developing nuclear energy for peaceful purposes? What is one argument against developing nuclear energy for peaceful purposes?

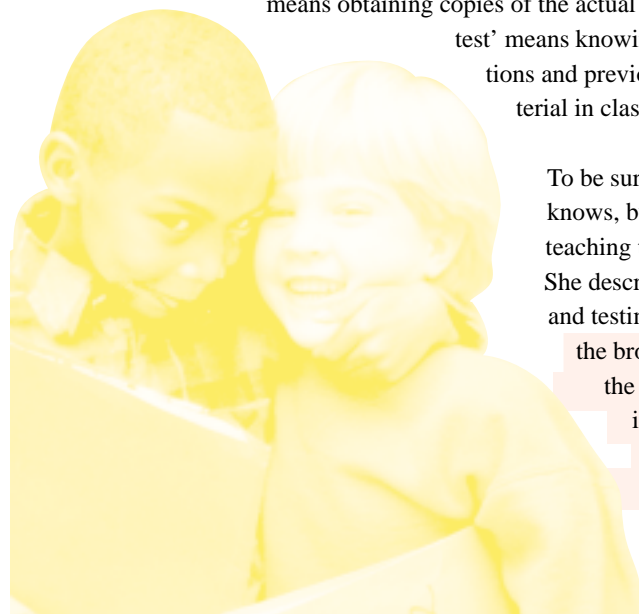
This question is taken from the National Assessment of Educational Progress (NAEP). Numerous other sample test questions for grade levels 4, 8, and 12 in several subject and skill areas at various difficulty levels are available on the "Sample Question Tool" on NAEP's Web site. (See Resources, page 4.)

Teaching to the Test

People often use the term "teaching to the test" to disparage testing, but some tests, such as Advanced Placement exams, set out a range of concepts and skills that students should be able to demonstrate after receiving good instruction.

Ruth Mitchell, principal partner at the Education Trust, makes an important distinction: "'Teaching the test' means obtaining copies of the actual test and teaching the answers. 'Teaching to the test' means knowing the content addressed from test specifications and previously administered tests and including that material in classroom work," she explains.

To be sure, tests can't measure everything a student knows, but to limit teaching to what is on the test—by teaching the test—shortchanges students, says Mitchell. She describes the proper relationship between teaching and testing: "Teaching should enable students to master the broad range of skills and knowledge required by the standards, and testing should drop a dipstick into the student's knowledge to diagnose how successful learning has been. There's nothing wrong with teaching to the test, if the test is worth teaching to."



One District's Experience With Accountability and Assessment

Over several years Texas experienced a range of state-wide educational reforms, "No Pass No Play" being one of the more publicized in this football-crazed state. Many of these had some level of impact, but the state needed a way to measure student achievement for all public school students.

The Texas Assessment of Academic Skills (TAAS) is a criterion-referenced test designed to measure a student's understanding of the Texas Essential Knowledge and Skills (TEKS), the standards for Texas public schools. It is administered in grades 3-8 and 10. As originally designed, TAAS was used to assess schools and there were no consequences or rewards for individual students until the 10th grade. If a student failed to pass the reading, math and writing sections of the 10th grade TAAS exit exam, he could continue to take the exam but could not graduate until passing TAAS.

Local Concern

The first TAAS in 1992 proved to be a wake up call for Waco. Before TAAS, in this district of 15,000-plus students, 80% of whom are identified as disadvantaged, the measurement for success had in effect been "my kid's happy and seems to be learning, so the schools must be good," according to Dr. Dot Snyder, school board member since 1997 for the Waco Independent School District (WISD). Although many of WISD's students graduated from high school ready for college or work, many others entered college unprepared to compete academically and in need of remediation classes. Employers had problems hiring graduates with basic skills.

Low scores on TAAS tests coupled with in-depth research of local data uncov-

ered a litany of problems in Waco's schools. A fundamental problem of promoting unprepared students to the next grade seemed to be the most significant.

Many of the students who failed TAAS had been promoted year after year. At one point over 50% of the students failed TAAS while only 2% were retained.

The cumulative effect of students promoted to the next grade without the needed skills became insurmountable. By high school many of these students dropped out or failed to get a diploma because they couldn't pass the TAAS exit exam.

New Policy and Support for Students

In the summer of 1997, the Waco school board implemented the most rigorous promotion policy in the state. Careful analysis of data guided the specifics of all decisions. The policy takes into account multiple criteria to determine when students are promoted to the next grade:

- district assessments and teacher/parent/principal input for grades 1 and 2;
 - an average TAAS score of 70 for grades 3 through 8 and the high school exit exam;
 - a 90% attendance rate each semester; and
 - passing grades in all classes.
- "The point of the policy," says Dr. Snyder, "is not to fail children, but to identify those students below grade level and intensely remediate them so that they can be successful learners through the rest of their school years."

Along with this higher level of accountability, the district provides significant early intervention, individualized instruction and remediation. To minimize the high stakes nature of this tough policy, students are given multiple opportunities to demonstrate their knowledge of the standards by retaking the TAAS test. The appeals process for attendance and grades violations is thorough and strictly enforced.

According to Dr. Snyder, "there is a difference between students who can't learn and students who won't learn." She believes when a student has no attendance or behavior problems and passing marks but does not pass TASS, the school needs to work harder to identify why and provide added support so the student is able to pass the test.

TAAS is designed to test state standards, and while the state standards are often described as "minimum skills," they are what Texas requires its students to learn. Parents expect teachers to teach these standards well. Consequently the district began to ensure that instruction focused on these standards. Schools re-ordered their curriculum to follow a district plan in this high mobility district, so children would be less likely to miss important information if they changed schools.

TAAS Data Are Used to Monitor Progress

WISD tests students every nine weeks with a version of TAAS that the state makes public. The data from these monitoring tests is broken down into specific learning objectives for each student and used to help teachers identify those who have fallen behind. Teachers are then able to work more intensely with them on specific skills. These tests also help teachers recognize where they may need to improve their teaching of the standards. Parents are notified and encouraged to participate in the remediation work. When students need extra time to master specific material, they are encouraged to attend tutorials and summer school.

Results

The first year of implementation saw 1350 students (19.3%) retained because of poor grades, excessive absences or low TAAS scores. However, because of careful data analysis prior to implementation, these numbers were not a surprise. Approximately \$1 million had been budgeted for remediation that first

year. Fewer dollars have had to be budgeted for remediation in subsequent years because annual retention rates have dropped by more than half since 1997.

Initially a number of students withdrew from the district to avoid the policy. When many of these same students returned to the public schools, sometimes further behind, the greater Waco community stepped up to meet the challenge of this policy with tutoring and mentoring programs that have grown in size and impact over time.

Now that students, teachers and parents have worked under this policy for a few years, people have come to expect improvement and are paying closer attention to the quality of instruction. The district recently hired a coordinator for staff development to assist teachers and principals in honing their professional skills and devising plans for professional growth.

Withstanding a lawsuit and accusations of racial discrimination, the district continues to fine-tune the policy to assure rigor and fairness. A small number of students face multiple retentions and need specific attention. Because the standards on which TAAS is based are "minimum level," higher local standards need to be further developed. And the meaningful inclusion of parents in the learning process remains weak at many schools.

Last year the American School Board Journal recognized the district's promotion policy with a Magna Award 2000. Most importantly, Waco's minority students have posted the greatest improvements of all student groups. The achievement gaps between student groups in Waco public schools have decreased at higher rates than the state averages. Though a small number of students face multiple retentions, scores for all groups of students in WISD continue to rise.

RESOURCES to a range of perspectives

(continued from back page)

cards can serve as a "yardstick" of what parents should expect. Their Web site (see "Supporting Documents"), explains how to read its report cards, offers a glossary, and provides a manual to explain how ratings are determined in various areas. <http://reportcard.ode.state.or.us>

Pittsburgh Council for Public Education—publishes a district report card available on the Public Education Network's Web site www.publiceducation.org. PCPE's phone: 412-434-0851

Public Education Fund in Providence, RI—group's process to engage parents in developing school progress reports is available in the Public Education Network's Standards Toolkit online at www.publiceducation.org/standards/tool/rptcard.htm PEF's phone: 401-454-1050

Access to the Tests and Test Items

The **Texas Education Agency** releases the Texas Assessment of Academic Skills (TAAS) and end-of-course tests annually. Schools can use them to gauge progress of students before they take the official TAAS test. Can help parents better understand the state standardized test, see examples of test items for several grade levels and subjects, and support advocacy to make test items public in other states. www.tea.state.tx.us/student.assessment/resources/release/index.html

Third International Mathematics and Science Study (TIMSS) Online Challenge, sponsored by the National Center for Research on Evaluation, Standards, and Student Testing (CRESST), provides sample items from the TIMSS to promote discussions of the TIMSS and math/science assessment items. <http://timssonline.cse.ucla.edu> TIMSS Web site: <http://timss.bc.edu>

National Center for Education Statistics (NCES)—annually releases portions of test items from the National Assessment of Educational Progress (NAEP); available on the Web in a "Sample Question Tool" with an extensive database searchable by grade level, subject, type of questions and/or difficulty. www.nces.ed.gov/nationsreportcard/ITMRLS/intro.shtml NAEP staff available to review testing booklets with small groups of parents. To request reviews, call NCES at 202-502-7300 or contact Dr. Peggy Carr by email at peggy_carr@ed.gov

Teaching should enable students to master the broad range of skills and knowledge required by the standards, and testing should drop a dipstick into the student's knowledge to diagnose how successful learning has been.

— Ruth Mitchell, Education Trust

Toolkit for Parents

According to Lynn Olson, senior editor at *Education Week*, there are four major components parents should examine in a state or district accountability system that holds schools responsible for performance:

- A report card on schools that includes information about student performance. Are these reports in parent-friendly language? Do they include the information that parents care about?
- A rating system for schools. Does this rating system indicate the level at which each school is performing and define acceptable levels of performance?
- Assistance provided for low-performing schools, in the form of extra resources for materials or professional development, or on-site technical assistance from the state department of education. Is the district participating in a state assistance program? If so, how? If not, is it eligible?
- Rewards for high, or significant gains in, performance and sanctions or consequences for consistent poor performance over time. What are the factors used to make decisions about whether schools or districts receive rewards for higher performance, or consequences for low performance?

Questions for Parents to Ask

The School District's Accountability System

- What are all the performance expectations for the schools in the district? What are schools, teachers and students being held accountable for? A score on a single test? Closing the achievement gap? For graduation rates? For demonstrating the knowledge and skills outlined in the state standards? How do parents and the public know whether schools and students are meeting the district's expectations?
- Is the district treating accountability as a help or a hammer? Is the information gathered from tests and other assessments of student performance used to identify students and schools that need extra help, and press for changes in professional development and how resources are allocated? Is the information used to punish schools or make a single decision about a student's future?

Tests

- What test is the state using to measure overall student performance? Is it a standards-based test? How can the district demonstrate to parents that the test matches the standards? Or is it a norm-referenced test?
- When is the test given? At key grade levels (e.g., 3rd, 8th, and 11th grades)? Every year? What month during the school year?
- What is the format of the test? Are all of the questions multiple choice, or are there also essay and/or short-answer questions? Do at least some of the questions require students to use higher order thinking skills to apply their knowledge?
- How can parents see examples of the test or of some test questions?

Test Preparation and Accommodations

- What are teachers doing to prepare students to do well on the test?
- What percentage of class time is spent taking practice tests?
- What kinds of accommodations are there for students with special needs (e.g., learning disabilities or non-English language speakers)?
- What happens if a student is absent on test day?
- What percentage of the district's students takes the test? Is there a large portion of students who don't take the test? Why? How might this affect the overall test scores? How might this affect other decisions based on the test data?

Reporting Results

- When are the scores reported? Do schools receive the scores in a reasonable amount of time after the test is taken?
- How are the test scores reported? Are they made available to the public in a school/district report card? Are the scores disaggregated, or broken down by race and other student and school characteristics?
- Do report cards include information on standards and expectations for all students?

Use of Results

- How are report cards being used by schools, parents and communities to improve student achievement?
- How are test results used? Are they used to diagnose needs for changes in teaching or curriculum, or for additional professional development in certain areas?
- What are the consequences—for students and schools—if students perform poorly on the test?
- Is this a "high-stakes" test? Is it the **only** factor used to make a decision about a student's academic future, or a school's progress? If so, do students have opportunities to retake the test if they do poorly the first time?
- Why might there be a discrepancy between a child's test score and his or her classroom grades?
- What other kinds of data are collected about student and teacher performance? Which of these data enables teachers to adjust lessons where students aren't learning at high levels?

Resources for Students and Teachers

- What extra learning help is available for students not meeting their grade level expectations or doing poorly on state tests (one-on-one help during the school day, summer school, mentoring, after-school, extra help one to two months before the state test)?
- How are assessments of what students are learning used to make decisions about programs and resources that help students learn or teachers teach? What data are used to help make decisions about staff development?

Resources from the State

- What forms of assistance does the state offer for low-performing schools?
- What resources is the district accessing from the state (in the form of technical assistance, school improvement grants, waivers)? Are there additional forms of assistance available?
- Is the district in a defensive posture about technical assistance from state?

RESOURCES to a range of perspectives

Accountability

Quality Counts—annual *Education Week* publication; contains valuable state-by-state data and rankings related to standards and accountability. 2001 issue, *A Better Balance: Standards and Accountability*, explores the status of standards implementation, standards-based assessment, and the variety of state accountability systems. 1999 issue, *Rewarding Results, Punishing Failure*, explores testing and interventions for low performing schools and contains a section on school/district report cards. Both available on *Education Week's* Web site or in hard copy from Editorial Projects in Education. 800-346-1834; www.edweek.org/sreports

Partnership for Learning—focused on communicating Washington State's school reform efforts; has developed several parent-friendly ways of explaining standards, accountability and testing. Their publications (see "Free Guides" and "Parents" sections of their Web site) can serve as models for parents and community groups seeking to explain local reform efforts. 800-550-5437; www.partnership-wa.org

Cross City Campaign for Urban School Reform—more information on "reciprocal accountability" model available on their Web site (under "Programs"). Also see Cross City's publication, "Beyond Finger-Pointing and Test Scores," examining interventions addressing low school performance. It offers helpful standards and recommendations for district-led interventions. Executive summary available on their Web site (under "Publications"). Entire report available for \$22. 312-322-4880; www.crosscity.org

Southern Regional Education Board—outlines important factors in making ratings and assistance work in Jim Watts' publication "Getting Results With Accountability: Rating Schools, Assisting Schools, Improving Schools." Good overview of the components of a comprehensive state system of accountability; includes helpful questions when seeking local information about each component. Available on SREB's Web site in Education Benchmarks 2000 Series Reports section (under "Recent Reports"). 404-875-9211; www.sreb.org

Education Trust—offers several publications and tools to help examine standards and data locally; staff members are knowledgeable about standards and accountability. This Washington D.C.-based group uses data deliberately to demonstrate achievement gaps and elements of successful high-poverty schools; Web site contains

interactive tool that compares achievement and resource levels with national indicators in each state. 202-293-1217; www.edtrust.org

The Testing Debate

Achieve, Inc.—non-partisan group started by governors and business leaders to promote standards-based reforms, tests and accountability. Published policy brief, "Testing: Setting the Record Straight," contains information about the evolution of present status of standards-based tests, and next steps to ensure that states develop rigorous tests that are aligned with standards. Available online, 202-624-1460; www.achieve.org

American Educational Research Association (AERA)—has position statement on high-stakes testing, based on the Standards for Educational and Psychological Testing from the National Research Council. Statement outlines conditions for appropriate use of tests. 202-223-9485; www.aera.net/about/policy/stakes/htm

National Center for Fair & Open Testing (FairTest)—watchdog group that promotes fair evaluation of students; known for their strict opposition to standardized testing and focus on inequities posed by standardized tests. Materials on testing practices and state-by-state analysis of statewide assessments available on the group's Web site. (617) 864-4810; www.fairtest.org

Wrightslaw—special education legal advocacy group posts information about Oregon's recent changes in its assessment system to better accommodate students with learning disabilities, and links to related resources. See "Legal News" section of their Web site. www.wrightslaw.com/lawlibrary_news.htm

Public Perceptions about Testing

Association for Supervision and Curriculum Development (ASCD)—provides survey of parent perceptions on testing, Q&A with testing experts, and test-taking tips for parents to use with their children. www.ascd.org/educationnews/pr/sylvan.html

Reports of School Progress

Maryland offers an extensive Web site on school improvement in the state, with resources to help understand the state test, good instruction, interventions and how to analyze and use data. www.md12.org

Oregon's efforts at school report (continued on page 5)