I Want to Know More
A *Leadership in Action* Supplement

I Want to Know More is a selection of information and resources for education leaders, parents, and community members who want to learn more about the teaching and learning strategies taking place in today's most innovative schools.

What Is Assessment?

When people hear the term "assessment," they most likely think: *standardized test*. But standardized tests are only one of many forms of assessment used in public education, each of which has its own specialized purpose.

Consider the dashboard on your car: when the check-engine light is on, you know that something might be wrong, but you don't know exactly *what*'s wrong. Conversely, even if the check-engine light is not on, there could still be something wrong with your car. Standardized tests are like check-engine lights: they can tell us broad information about how our schools are functioning and our students are performing (for example, that eighth-grade math scores are low), but they don't tell *why* the scores are low or high, what's specifically working or not working in schools.

When it comes to assessment, teachers, students, and policy makers need far more than check-engine lights—they need an entire dashboard of information about the learning process to determine precisely what needs to be adjusted or fixed to meet every learning goal.

Defining Assessment

The term assessment refers to the wide variety of methods or tools that educators use to evaluate, measure, and document the academic readiness, learning progress, skill acquisition, or educational needs of students.

While assessments are often equated with traditional tests—especially the standardized tests developed by testing companies and administered to large populations of students—educators use a diverse array of assessment tools and methods to measure everything from a four-year-old's readiness for kindergarten to a twelfth-grade student's comprehension of advanced physics. Just as academic lessons have different functions, assessments are typically designed to measure specific elements of learning—e.g., the level of knowledge a student already has about the concept or skill the teacher is planning to teach or the ability to comprehend and analyze different types of texts and readings. Assessments also are used to identify individual student weaknesses and strengths so that educators can provide specialized academic support, educational programming, or social services. In addition, assessments are developed by a wide array of groups and individuals, including teachers, district administrators, universities, private companies, state departments of education, and groups that include a combination of these individuals and institutions.

While assessment can take a wide variety of forms in education, the following descriptions provide a representative overview of a few major forms of educational assessment.

Assessments are used for a wide variety of purposes in schools and education systems:

- **High-stakes assessments** are typically standardized tests used for the purposes of accountability—i.e., any attempt by federal, state, or local government agencies to ensure that students are enrolled in effective schools and being taught by effective teachers. In general, "high stakes" means that important decisions about students, teachers, schools, or districts are based on the scores students

achieve on a high-stakes test, and either punishments (sanctions, penalties, reduced funding, negative publicity, not being promoted to the next grade, not being allowed to graduate) or accolades (awards, public celebration, positive publicity, bonuses, grade promotion, diplomas) result from those scores. For a more detailed discussion, see high-stakes test.

- **Pre-assessments** are administered before students begin a lesson, unit, course, or academic program. Students are not necessarily expected to know most, or even any, of the material evaluated by pre-assessments—they are generally used to (1) establish a baseline against which educators measure learning progress over the duration of a program, course, or instructional period, or (2) determine general academic readiness for a course, program, grade level, or new academic program that student may be transferring into.

- **Formative assessments** are in-process evaluations of student learning that are typically administered multiple times during a unit, course, or academic program. The general purpose of formative assessment is to give educators in-process feedback about what students are learning or not learning so that instructional approaches, teaching materials, and academic support can be modified accordingly. Formative assessments are usually not scored or graded, and they may take a variety of forms, from more formal quizzes and assignments to informal questioning techniques and in-class discussions with students.

- **Summative assessments** are used to evaluate student learning at the conclusion of a specific instructional period—typically at the end of a unit, course, semester, program, or school year. Summative assessments are typically scored and graded tests, assignments, or projects that are used to determine whether students have learned what they were expected to learn during the defined instructional period.

> While Formative assessments are commonly said to be for learning because educators use the results to modify and improve teaching techniques during an instructional period, while summative assessments are said to be of learning because they evaluate academic achievement at the conclusion of an instructional period. Or as assessment expert Paul Black put it, "When the cook tastes the soup, that's formative assessment. When the customer tastes the soup, that's summative assessment."

- **Interim assessments** are used to evaluate where students are in their learning progress and determine whether they are on track to performing well on future assessments, such as standardized tests, end-of-course exams, and other forms of "summative" assessment. Interim assessments are usually administered periodically during a course or school year (for example, every six or eight weeks) and separately from the process of instructing students (i.e., unlike formative assessments, which are integrated into the instructional process).

- **Placement assessments** are used to "place" students into a course, course level, or academic program. For example, an assessment may be used to determine whether a student is ready for Algebra I or a higher-level algebra course, such as an honors-level course. For this reason, placement assessments are administered before a course or program begins, and the basic intent is to match students with appropriate learning experiences that address their distinct learning needs.

- **Screening assessments** are used to determine whether students may need specialized assistance or services, or whether they are ready to begin a course, grade level, or academic program. Screening assessments may take a wide variety of forms in educational settings, and they may be developmental, physical, cognitive, or academic. A preschool screening test, for example, may be used to determine whether a young child is physically, emotionally, socially, and intellectually ready to begin preschool, while other screening tests may be used to evaluate health, potential learning disabilities, and other student attributes.

Assessments are also designed in a variety of ways for different purposes:

- **Standardized assessments** are designed, administered, and scored in a standard, or consistent, manner. They often use a multiple-choice format, though some include open-ended, short-answer questions. Historically, standardized tests featured rows of ovals that students filled in with a number-two pencil, but increasingly the tests are computer-based. Standardized tests can be administered to large student populations of the same age or grade level in a state, region, or country, and results can be compared across individuals and groups of students. For a more detailed discussion, see standardized test.

- **Standards-referenced or standards-based assessments** are designed to measure how well students have mastered the specific knowledge and skills described in local, state, or national learning standards. Standardized tests and high-stakes tests may or may not be based on specific learning standards, and individual schools and teachers may develop their own standards-referenced or standards-based assessments. For a more detailed discussion, see proficiency-based learning.

- **Common assessments** are used in a school or district to ensure that all teachers are evaluating student performance in a more consistent, reliable, and effective manner. Common assessments are used to encourage greater consistency in teaching and assessment among teachers who are responsible for teaching the same content, e.g. within a grade level, department, or content area. They allow educators to compare performance results across multiple classrooms, courses, schools, and/or learning experiences (which is not possible when educators teach different material and individually develop their own distinct assessments). Common assessments share the same format and are administered in consistent ways—e.g., teachers give students the same instructions and the same amount of time to complete the assessment, or they use the same scoring guides to interpret results. Common assessments may be "formative" or "summative." For more detailed discussions, see coherent curriculum and rubric.

- **Performance assessments** typically require students to complete a complex task, such as a writing assignment, science experiment, speech, presentation, performance, or long-term project, for example. Educators will often use collaboratively developed common assessments, scoring guides, rubrics, and other methods to evaluate whether the work produced by students shows that they have learned what they were expected to learn. Performance assessments may also be called "authentic assessments," since they are considered by some educators to be more accurate and meaningful evaluations of learning achievement than traditional tests. For more detailed discussions, see authentic learning, demonstration of learning, and exhibition.

- **Portfolio-based assessments** are collections of academic work—for example, assignments, lab results, writing samples, speeches, student-created films, or art projects—that are compiled by students and assessed by teachers in consistent ways. Portfolio-based assessments are often used to evaluate a "body of knowledge"—i.e., the acquisition of diverse knowledge and skills over a period of time. Portfolio materials can be collected in physical or digital formats, and they are often evaluated to determine whether students have met required learning standards. For a more detailed discussion, see portfolio.

The purpose of an assessment generally drives the way it is designed, and there are many ways in which assessments can be used. A standardized assessment can be a high-stakes assessment, for example, but so can other forms of assessment that are not standardized tests. A portfolio of student work can be a used as both a "formative" and "summative" form of assessment. Teacher-created assessments, which may also be created by teams of teachers, are commonly used in a single course or grade level in a school, and these assessments are almost never "high-stakes." Screening assessments may be produced by universities that have conducted research on a specific area of child development, such as the skills and attributes that a student should have when entering kindergarten to increase the likelihood that he or she will be successful, or the pattern of behaviors, strengths, and challenges that suggest a child has a particular learning disability. In short, assessments are usually created for highly specialized purposes.

Assessment and School Improvement

While educational assessments and tests have been around since the days of the one-room schoolhouse, they have increasingly assumed a central role in efforts to improve the effectiveness of public schools and teaching. Standardized-test scores, for example, are arguably the dominant measure of educational achievement in the United States, and they are also the most commonly reported indicator of school, teacher, and school-system performance.

As schools become increasingly equipped with computers, tablets, and wireless internet access, a growing proportion of the assessments now administered in schools are either computer-based or online assessments—though paper-based tests and assessments are still common and widely used in schools. New technologies and software applications are also changing the nature and use of assessments in innumerable ways, given that digital-assessment systems typically offer an array of features that traditional paper-based tests and assignments cannot. For example, online-assessment systems may allow students to log in and take assessments during out-of-class time or they may make performance results available to students and teachers immediately after an assessment has been completed (historically, it might have taken hours, days, or weeks for teachers to review, score, and grade all assessments for a class). In addition, digital and online assessments typically include features, or "analytics," that give educators more detailed information about student performance. For example, teachers may be able to see how long it took students to answer particular questions or how many times a student failed to answer a question correctly before getting the right answer. Many advocates of digital and online assessments tend to argue that such systems, if used properly, could help teachers "personalize" instruction—because many digital and online systems can provide far more detailed information about the academic performance of students, educators can use this information to modify educational programs, learning experiences, instructional approaches, and academic-support strategies in ways that address the distinct learning needs, interests, aspirations, or cultural backgrounds of individual students. In addition, many large-scale standardized tests are now administered online, though states typically allow students to take paper-based tests if computers are unavailable, if students prefer the paper-based option, or if students don't have the technological skills and literacy required to perform well on an online assessment.

Given that assessments come in so many forms and serve so many diverse functions, a thorough discussion of the purpose and use of assessments could fill a lengthy book. The following descriptions, however, provide a brief, illustrative overview of a few of the major ways in which assessments—especially assessment results—are used in an attempt to improve schools and teaching:

- **System and school accountability**: Assessments, particularly standardized tests, have played an increasingly central role in efforts to hold schools, districts, and state public-school systems "accountable" for improving the academic achievement of students. The most widely discussed and far-reaching example, the 2001 federal law commonly known as the No Child Left Behind

Act, strengthened federal expectations from the 1990s and required each state develop learning standards to govern what teachers should teach and students should learn. Under No Child Left Behind, standards are required in every grade level and content area from kindergarten through high school. The law also requires that students be tested annually in grades 3-8 and at least once in grades 10-12 in reading and mathematics. Since the law's passage, standardized tests have been developed and implemented to measure how well students were meeting the standards, and scores have been reported publicly by state departments of education. The law also required that test results be tracked and reported separately for different "subgroups" of students, such as minority students, students from low-income households, students with special needs, and students with limited proficiency in English. By publicly reporting the test scores achieved by different schools and student groups, and by tying those scores to penalties and funding, the law has aimed to close achievement gaps and improve schools that were deemed to be underperforming. While the No Child Left Behind Act is one of the most controversial and contentious educational policies in recent history, and the technicalities of the legislation are highly complex, it is one example of how assessment results are being used as an accountability measure.

- **Teacher evaluation and compensation**: In recent years, a growing number of elected officials, policy makers, and education reformers have argued that the best way to improve educational results is to ensure that students have effective teachers, and that one way to ensure effective teaching is to evaluate and compensate educators, at least in part, based on the test scores their students achieve. By basing a teacher's income and job security on assessment results, the reasoning goes, administrators can identify and reward high-performing teachers or take steps to either help low-performing teachers improve or remove them from schools. Growing political pressure, coupled with the promise of federal grants, prompted many states to begin using student test results in teacher evaluations. This controversial and highly contentious reform strategy generally requires fairly complicated statistical techniques—known as value-added measures or growth measures—to determine how much of a positive or negative effect individual teachers have on the academic achievement of their students, based primarily on student assessment results.

- **Instructional improvement**: Assessment results are often used as a mechanism for improving instructional quality and student achievement. Because assessments are designed to measure the acquisition of specific knowledge or skills, the design of an assessment can determine or influence what gets taught in the classroom ("teaching to the test" is a common, and often derogatory, phrase used to describe this general phenomenon). Formative assessments, for example, give teachers in-process feedback on student learning, which can help them make instructional adjustments during the teaching process, instead of having to wait until the end of a unit or course to find out how well students are learning the material. Other forms of assessment, such as standards-based assessments or common assessments, encourage educators to teach similar material and evaluate student performance in more consistent, reliable, or comparable ways.

- **Learning-needs identification**: Educators use a wide range of assessments and assessment methods to identify specific student learning needs, diagnose learning disabilities (such as autism, dyslexia, or nonverbal learning disabilities), evaluate language ability, or determine eligibility for specialized educational services. In recent years, the early identification of specialized learning needs and disabilities, and the proactive provision of educational support services to students, has been a major focus of numerous educational reform strategies. For a related discussion, see academic support.

LEADERSHIP IN ACTION
is a new england secondary school consortium resource
newenglandssc.org/leadership_in_action

5