

CHAPTER ONE

Achieving Equity Through Standards and Assessments

PREPARING ALL STUDENTS FOR SUCCESS

Holding all students to high academic standards is the centerpiece of a national agenda to improve schools and ensure that no child is left behind in the journey toward the American dream. The evolution of standards over the past decade has been driven by the need to define what all students should learn in school in order to participate successfully in the twenty-first century. It has also been driven by widespread recognition that today's world requires many skills that are not being taught in schools. The issue is not whether schools are better or worse than they used to be, but whether public schools are preparing all children to succeed in today's world. Today's workplace already requires that individuals understand multidimensional problems, design solutions, plan their own tasks, evaluate results, and work cooperatively with others. These requirements demand that

2 STANDARDS-BASED INSTRUCTION AND ASSESSMENT

American students learn a new set of competencies and foundation skills, and they demand an approach to education that is very different from the education system that served the industrial age (Lachat, 1994).

Many studies and reports have concluded that unless educational performance in the United States improves dramatically, American workers will not be able to use the technologies that will create most of the world's jobs and economic growth in the twenty-first century. The impetus for educational reform has been driven by other sources as well: cognitive research on the processes of learning; recognition of the disparities in educational opportunity for diverse student groups; concerns about the quality of the teaching force; and criticisms of the bureaucratic and departmentalized nature of the education system (Berman et al., 1995).

Historically, the standards movement reflected the belief that it is in the national interest to educate all students to their full potential. This belief became law in 1994 when Congress legislated changes in the Elementary and Secondary Education Act (ESEA), which raised the bar of achievement for every American student. This federal legislation required states to replace minimum standards for poor and academically disadvantaged students with challenging standards for all students, to develop new accountability systems aligned with the state standards, and to hold all students to the same performance standards (Riddle, 1999). The act was strengthened by the 2001 No Child Left Behind (NCLB) legislation, which defined stronger accountability mandates and is providing added momentum to the emphasis on higher standards for all students. Standards are thus the core of a federally mandated accountability system directly aimed at improving the quality of teaching and learning in American schools. The belief at the federal level is that American education will significantly change for the better if it sets high standards and uses standards-based assessments to hold districts, schools, and teachers accountable for student learning. Touching on every aspect of the education system, reform

mandates are challenging long-held assumptions about how education should be conducted, particularly for students who have had the least access to high quality learning environments. The expectation is that schools will educate *all* children to be effective thinkers, problem solvers, and communicators.

The Standards Model

The standards model is based on several important assumptions: that educators can define standards for what is most important for students to know and be able to do in today's society; that most students will be able to achieve the standards; that student performance may differ in demonstrating proficiency but will still reflect the defined standards; and that standards will allow for fair and consistent assessment of diverse student performances (Taylor, 1994). Two types of standards provide the foundation for standards-based curriculum, instruction, and assessment: content standards and performance standards.

Content Standards: These standards define what children should know and be able to do. They describe the essential knowledge, skills, and understandings necessary to attain proficiency in a content area. Content standards define the focus of curriculum and instruction within and across content areas, and they describe what teachers are expected to teach and what students are expected to learn. They represent the curricular priorities of the state or district, and they can serve as starting points for curriculum improvement because they describe what is important for all students to learn in the various subject areas. *Benchmarks* are subcomponents of content standards—they identify the expected understandings and skills for a content standard at different grade levels.

Performance Standards: These standards identify levels of performance for the knowledge, skills, and understandings defined in the content standards. They set specific expectations

4 STANDARDS-BASED INSTRUCTION AND ASSESSMENT

for various levels of proficiency and provide explicit definitions and concrete examples of what students must demonstrate in order to be considered proficient in the knowledge and skills of the content standards. The emphasis is on the types of evidence that will verify the degree of proficiency the student has demonstrated. Performance standards are defined in terms of various levels of proficiency.

The standards movement has challenged educators to no longer accept low expectations for so many students. By creating stronger connections between what and how students learn in school and how they will be expected to perform in adult life, the standards movement demands instructional approaches that develop students' reasoning and problem-solving skills. This new vision of teaching and learning requires educators and the public to understand that high standards are as important in education as they are in the medical profession, in licensing pilots, or in international sports competitions such as the Olympics. They define what is essential for successful performance and encourage people to strive for the best.

Across the country, state education agencies have been implementing the standards model to power a nationwide curriculum reform movement. As of the fall of 2003, all 50 states plus the District of Columbia and Puerto Rico had completed and implemented K–12 content standards in the core subjects of English/language arts and mathematics, and 48 states had completed and implemented standards for science and social studies/history. The intent of the state content standards is that their use will help administrators, teachers, and parents develop common understandings about what students should learn.

Developing high quality education standards is a complex process, however, as was shown by the national efforts to develop content standards during the 1990s. Likewise, the process of standards development at the state level has highlighted that not all standards are equal, and the quality of

state standards has come under critical review by different organizations (American Federation of Teachers, 1999; Council for Basic Education, 1998; Finn & Petrelli, 2000). Although the different organizations conducting the reviews gave very different ratings of the standards, the consensus was that state standards have improved over the past few years in the extent to which they are specific and measurable; in most states, however, they still need to be improved (Kendall & Marzano, 2000).

STANDARDS AND EDUCATIONAL EQUITY

The most pressing issue of education reform in the twenty-first century is how to ensure a high quality education for all of the nation's children, an issue that is growing in proportion to the rapidly changing demographics of America's schools. The focus on standards in federal and state mandates has shifted the emphasis from "access for all" to "high quality learning for all." Beneath the surface of standards-based reform is the question of whether the American dream truly belongs to all students and whether American society is morally committed to equal educational opportunity. In her study of the development of standards in America from the mid-1980s to the mid-1990s, Ravitch (1995) highlighted that standards-based educational reform reflected a unique convergence of people committed to excellence and people committed to quality. Her point was that people who worried most about excellence looked to standards to raise achievement; people who worried most about equality looked to standards to provide students with equal access to challenging curricula and learning experiences. Together they forged an unusual and effective alliance. From an equity perspective, the emphasis on high standards for all students springs from a vision of society where people of different backgrounds, cultures, and perceived abilities have equal access to a high quality education.

6 STANDARDS-BASED INSTRUCTION AND ASSESSMENT

The standards movement reflects many dimensions of a new society that has emerged over the past two decades, creating totally new equity demands for schools. We are seeing an unprecedented commitment to educate *all* students to be effective thinkers, problem solvers, and communicators so they can participate as productive members of the global community. As waves of immigration bring people from all over the world to our nation, we are becoming the most ethnically, culturally, and linguistically diverse society that has ever existed. In combination, these conditions and commitments are creating a new order for schools. Never before have schools been asked to ensure that all students achieve publicly defined standards of learning. Never before have we asked schools to consider higher-order skills as core skills to be acquired by all students. Never before have teachers been faced with such diversity (Lachat, 1999a).

Historically, our education system was organized to develop basic skills for all students and higher-order proficiencies for those who were college bound. The current reform agenda disclaims this two-tier system and challenges schools to demonstrate a core belief of public education in America—that all children can learn at high levels, given the time, tools, teaching, and encouragement to do so. However, this shift in thinking is coming at a time when schools are faced with more diversity in student populations than they have ever experienced. As standards become the foundation for curriculum and instruction, many school administrators and teachers are uncertain about how to use those standards with diverse populations. Some of the questions focus on equity issues and how the same high standards can be used for all students, while other questions focus on their use with students who have not developed full proficiency in English. Nevertheless, what the standards movement will accomplish for culturally diverse students is to leave no doubt that schools must engage these students in higher-order instruction.

Evidence is growing that the use of public standards as the basis for school accountability has already motivated school

administrators to avoid tracking into limiting groups those students who are not fully proficient in English, and to determine how these students can develop essential knowledge and skills. Given today's increasingly diverse student populations, the question is no longer whether it is feasible to provide a high quality education for students who vary widely in their characteristics, learning styles, levels of English proficiency, and educational needs. The question is how to reform curriculum and instruction and improve teachers' abilities to respond to diversity so that high quality learning becomes the norm for all students.

Equity and Opportunity to Learn

From an equity perspective, setting high standards for all students means that we must believe that the quality of education offered to "the best and the brightest" should be the quality of education available to all. "Watered down" curricula that deny students adequate preparation for success in an increasingly demanding world are unacceptable, and opportunity to learn has taken on a new meaning in terms of the curricular offerings and instructional methods that should be available to all students. But whether education reforms prove helpful to student populations that have been denied access to a challenging, high quality education will depend on whether these students gain access to the essential resources and conditions that support learning and achievement. Education standards and accountability mandates will not improve student learning unless they are accompanied by policies and practices that directly address inequities in conditions for learning. Curriculum reform and professional development for teachers are not enough to close achievement gaps for students who attend schools located in high-poverty environments. A vast number of students simply do not have equal access to the quality of education necessary to achieve high standards of learning.

8 STANDARDS-BASED INSTRUCTION AND ASSESSMENT

The problem with assuming education standards will improve teaching practices for poor and language minority groups is that this assumption ignores the grossly inadequate conditions that exist in the schools they attend (Darling-Hammond, 1997; Winfield, 1995). Many of these students are in schools that receive inadequate funding, have inadequate instructional materials and technology resources, and have difficulty recruiting qualified teachers. Students in these schools have minimal opportunities to develop the proficiencies reflected in emerging standards and new assessments. At the classroom level, their opportunity to learn is influenced by curriculum content, teacher beliefs, the quality of instruction, time spent on academic tasks, the nature of teacher-student interactions, and the feedback and incentives provided to them (Neill, 1995). In addition, family support, school safety, and school climate also influence the learning of these students.

Gordon (1992) has long cautioned that education reform should not occur in a vacuum but must consider the complex societal conditions that control access to essential resources:

[T]here are those of us who are sympathetic to standards and assessment, but insist that it is immoral to begin by measuring outcomes before we have seriously engaged the equitable and sufficient distribution of inputs, that is, opportunities and resources essential to the development of intellect and competence. (p. 2)

Widespread concerns exist about equity in the availability of school resources essential to implementing standards-based reform. Several components of organizational capacity have been identified:

- School administrators must provide the organizational leadership to implement the multiple elements of education reform.

Achieving Equity Through Standards and Assessments 9

- Schools must have certified and qualified personnel.
- Teachers must have the knowledge, skills, and capabilities to provide high quality instruction to diverse student populations.
- Financial and programmatic resources that include the high quality instructional materials, laboratories, libraries, computer facilities, and technology needed to support standards-based learning must be available.
- Teachers and administrators must have access to high quality professional development.
- The school environment must be safe and secure.
- Schools must have data system and reporting capabilities to respond to local, state, and federal requirements to demonstrate progress and accountability. (Corcoran & Goertz, 1995; Darling-Hammond, 1997; Lachat, 1999b; Newmann, King, & Rigdon, 1997; Ruiz de Velasco & Fix, 2000)

If all students are to have an equal opportunity to learn the knowledge and skills essential to success, inequities in how educational resources are allocated must be eliminated. The school capacity (i.e., a school's ability to respond to multiple demands) and resource implications of "high standards for all students" have yet to be fully determined, particularly for schools that serve poor and linguistically diverse student populations. But the mandates driving standards-based education reform can provide the much-needed leverage for addressing equity issues—the leverage to ensure all schools and teachers are aiming at the same high goals; the leverage to reduce inequities in the resources available to schools; and the leverage to ensure all students have equal access to instructional environments that support learning and achievement. Organizing school curricula around clearly defined standards may put an end to the inferior education that now deprives many children of the chance to study a challenging curriculum. However, the success of standards-based reform is directly connected to a comprehensive set of

10 STANDARDS-BASED INSTRUCTION AND ASSESSMENT

changes involving access to resources, access to highly skilled teachers, access to high quality instruction, and a safe and supportive school environment.

THE CULTURE OF STANDARDS-BASED ASSESSMENT

States have adopted the standards model to develop the assessment systems that will be used to measure the progress of all students. These emerging systems represent a different way of thinking about large-scale assessment, and they are playing a central role in school reform. For the first time, student learning is being measured against publicly defined standards, and performance-based assessment methods are being used to measure student proficiencies. These assessments focus on improving student learning and as a result are creating new concepts of accountability for schools. Standards-based assessments have several important features:

- They focus attention on what is most important to learn.
- They compare students to a standard of proficiency, not to other students.
- They are linked to curriculum and instruction.
- They are intended to establish accountability, as well as stimulate improvement.

Standards-based assessments represent a profound and sweeping shift away from the beliefs of a testing culture to the beliefs of an assessment culture, and they reflect changing assumptions about the nature of intelligence and how people learn. Testing and assessment cultures have radically different belief systems and goals. Helping educators and the public understand the implications of this change in belief systems is one of the challenges of education reform.

The Traditional Testing Culture

Traditional testing practices in America did not match the vision of standards-based reform. Historically, standardized tests had not been designed to measure complex skills and performance abilities and, as a result, often drove instruction toward lower-order cognitive skills (Darling-Hammond, 1997; Linn, 2000). Therefore, testing practices often contributed to the educational problems plaguing many schools. Based on a measurement model, the testing culture that dominated American education for most of the twentieth century assumed that intelligence and learning capacity were fixed traits that could be predicted. Because of this assumption, educators believed that students had an inherent level of intelligence that governed what they were able to learn. Therefore, the aim of testing was to rank students for purposes of comparison and placement. Testing was traditionally used to sort students according to their abilities (presumed to be inherent) and then to track them into “appropriate” educational programs (Taylor, 1994). Testing practices thus served to limit the access that some students had to higher-level learning opportunities, and students from low-income and varied cultural and language backgrounds particularly experienced the negative impacts of testing policies.

The impact of ranking on American education was far reaching. Founded on early-twentieth-century theories that treated intelligence as a unitary, fixed trait, America’s testing culture encouraged the belief that individuals could be ranked according to mental capacities. Scores representing children’s abilities were positioned relative to one another on a normal curve, causing educators to confuse a student’s rank with his or her potential for success (Wolf, Bixby, Glenn, & Gardner, 1991). The use of the normal curve as the dominant profile for showing student achievement led to widespread acceptance among policymakers, administrators, and teachers that a significant percentage of students would fail. Therefore, the

12 STANDARDS-BASED INSTRUCTION AND ASSESSMENT

belief system of the testing culture provided the public school system with a scientific rationale for tracking, and schools were not held accountable for the learning and academic success of all students. In short, the quality of education made available to many students was undermined by the testing policies and practices used to define and monitor their learning, and they were denied the opportunity to develop the capacities they needed to succeed.

The testing culture also did not emphasize complex and rich ways of demonstrating learning. Focusing on a narrow range of cognitive abilities that magnified differences among students, the testing culture valued accuracy, speed, and easily quantifiable skills. The selection of test items, which ranged from very easy to very difficult, was based on how well items discriminated between high and low scores so that scores could be easily differentiated and ranked (Farr & Trumbull, 1997). High proportions of those students who received low scores were poor students and students learning English as a second language. Because of their low test scores, these students were placed in low-level classes. Until the mid-1990s, testing was predominantly used to group these students according to levels of ability and provided a long-standing excuse for the limited learning opportunities that were provided to so many students.

Today's Assessment Culture

The focus on high standards for all students that began in the 1990s also caused a re-evaluation of the purpose of assessment by shifting the emphasis away from ranking students against test norms to an emphasis on improving student learning. This shift in emphasis was closely tied to the recognition that traditional testing practices had fueled inequities in education by relegating many students to a low-level education that limited their learning opportunities and life choices. The emerging assessment culture uses assessment as a tool to help schools understand their

students' proficiencies against publicly defined learning standards. The use of assessment has therefore changed from deciding which students will have access to a high quality education to determining whether all students are learning at high levels. Under this form of assessment, students are compared to a standard of performance, not to other students. At the policy level, this shift is significant. When students are measured against publicly defined standards of achievement, rather than against national norms established by test companies, more open discussion of the appropriateness of the assessment standards is possible. If the conception, development, and interpretation of assessment become open processes, then hidden biases can become more visible and more of the public will have a clear sense of what counts in our schools (Garcia & Pearson, 1994).

New forms of assessment are based on cognitive research that created new understandings about how children learn. The research challenged outmoded theories of learning that fostered the testing of sequential rote instruction but not critical thinking. It created a new vision of assessment that was developmental, went beyond a narrow range of skills, and offered new ways of examining how students think and perform when solving complex problems. Underlying the new approach is the belief that intelligence is not a fixed trait; instead, learning potential is considered to be developmental and a function of experience. An assessment culture recognizes that intelligence is multifaceted—that people's multiple intelligences have varying degrees of strength, are at various stages of development, and thus cannot be accurately ranked according to a single dimension (Gardner, 1993).

By shifting from a "measurement model" to a "standards model," assessment now focuses on how student performance develops relative to standards of excellence, not on how each student ranks against other students. This profoundly changes the playing field for students who are not yet fully proficient in English and who must have access to

14 STANDARDS-BASED INSTRUCTION AND ASSESSMENT

educational opportunities that foster development learning. The onus is now on the schools to provide these opportunities rather than sorting students into limited opportunities. A summary of the differences between the current assessment culture and the testing culture that dominated American education for most of the twentieth century is shown in Figure 1.1.

Figure 1.1 Comparison of a testing culture and an assessment culture

The Testing Culture	The Assessment Culture
Views intelligence as a unitary, fixed trait.	Views intelligence as multifaceted and learning potential as developmental.
Based on a measurement model that treats students' abilities as relative positions on a normal curve.	Based on a standards model where achievement is criterion-referenced.
Emphasizes accuracy, speed, and easily quantifiable skills.	Focuses on what students can do (performance), not just what they know (content domain).
Focuses on a narrow range of cognitive abilities.	Emphasizes complex and rich ways of demonstrating learning.
Views testing and instruction as separate activities.	Regards assessment as central to instruction.
Determines how students rank and compare with each other.	Determines how students perform relative to standards of excellence.
Uses test results to sort students into classes and courses emphasizing differential abilities.	Uses assessment results to improve teaching and learning.

The Use of Performance Assessments

When assessments are tied to standards, students must demonstrate what they know and can do through a range of “performances,” and more emphasis is placed on tasks that involve higher-order thinking and more complex problem solving. Performance assessments offer a better way of measuring the attainment of high learning standards than do traditional assessments. Wiggins (1989) was an early proponent of performance assessment as a more appropriate and meaningful way to assess student learning, suggesting that student performances that were “authentic” to the concepts, knowledge, and skills of a discipline and based on real-world problems could be identified for all subject areas. When he first recommended that these identified performances should form the foundation of new assessment programs, his writings provoked strong response from policymakers and educators. At the time, although many educators were dissatisfied with standardized achievement tests, they also saw traditional tests as the only way to ensure fair and reliable large-scale testing.

Today, however, performance-based assessment is rapidly becoming accepted as a promising vehicle for state assessment programs that are showing an increased emphasis on performance approaches. Most of the states are using some form of performance assessment in their assessment programs, particularly for the content areas of writing, mathematics, and reading. The challenge of large-scale assessment is that performance assessments rely on the judgment of those scoring the tests. The assessor must apply clearly articulated performance criteria in making a professional judgment about the level of proficiency demonstrated. The scoring process has to look beyond right or wrong answers; it also has to consider the thoughtfulness of the procedures used to carry out the task or solve the problem (Baker, 1997).

New Criteria for Validity and Reliability

The emergence of standards-based state assessment systems has forced experts who deal with the technical aspects of

16 STANDARDS-BASED INSTRUCTION AND ASSESSMENT

assessment to rethink their methods for assuring quality. The litmus test for any measuring instrument has always been its degrees of reliability (the degree to which the test yields the same results on repeated trials) and validity (the degree to which a test measures what it is intended to measure and to which inferences made based on a test's results are appropriate and useful for all students). In the past, test development often sacrificed validity to achieve reliability, in effect sacrificing the students' interests to the interests of test developers (Baker & Linn, 2002). However, new approaches to assessment have led some to question the role that reliability has traditionally played in assessment. Content validity, or the ability to understand what student performance reveals about learning, is of primary importance in standards-based assessment. Because of this, the traditional emphasis on reliability is being re-examined.

By their nature, performance assessments often require integrated knowledge and skills; are far less standardized than traditional tests; and allow for more latitude in design, in student response, and in scorer interpretation. Because of this, establishing reliability has been a major issue in new state assessment systems. Establishing consistency in scoring among well-trained raters has been more successful than establishing consistency across tasks. For example, test developers have had difficulty in establishing acceptable levels of comparability (reliability) across tasks intended to address the same skills. Some research shows that consistency of performance across tasks is influenced by the extent to which tasks share comparable features and also reflect the types of instruction students had received. Research also has shown that variations in task performance may be attributable to differences in students' prior knowledge and their experiences in performing similar tasks (Pellegrino, Chudowsky, & Glaser, 2001). This has significant implications for students who are not from the mainstream culture and who are not fully proficient in English.

Researchers have given increasing attention to the validity criteria that should characterize the use of performance

assessments in large-scale statewide assessment programs (Baker & Linn, 2002). Leading assessment specialists recommend that these assessments exemplify current content standards for what students should know and be able to do in various subject areas, and they should also contain explicit standards for rating or judging performance. The assessments should require that complex cognition be demonstrated through knowledge representation and problem solving. The validity criteria recommended by assessment specialists also stress that performance assessments be fair to students of different backgrounds and meaningful to students and teachers, incorporating competencies that can be taught and learned.

The emergence of large-scale performance assessments has therefore highlighted the importance of establishing a connection between validity standards and the interpretation and use of assessment results for diverse student populations. Due to the growing intolerance of test policies that limit students' access to learning, the concept of *consequential validity* has emerged. Consequential validity stresses that an assessment's *use* is what matters—that is, whether the use of assessment results produces positive consequences for students and for the teaching and learning process (Farr & Trumbull, 1997). It draws attention to the inequities produced when test results are used to limit educational opportunities. Consequential validity emphasizes that the use of assessment results is as important as technical concerns about reliability and content validity, and that tests must be evaluated in terms of their effects on the lives of students.

Accountability and Equity

The premise of high standards for all students is the greatest accountability challenge of the standards movement, particularly in low-performing schools with high proportions of culturally and linguistically diverse students. For the first time, schools are expected to ensure that *all* students achieve publicly defined standards of learning. For the first time,

18 STANDARDS-BASED INSTRUCTION AND ASSESSMENT

schools must consider “higher-order” skills as *core* skills to be acquired by all students, not just the most gifted. This represents a new way of thinking, a paradigm shift calling for high expectations for every student in every school, not just some students in some schools. This shift represents a very different mission for schools and a new emphasis on accountability for the success of all students.

Because standards-based assessments are part of the push toward higher levels of learning, they drive demands that schools verify that all students, including students who are not fully proficient in English, are achieving at acceptable levels. When policymakers link higher standards of performance to school accountability, however, they provoke considerable discussion and debate. To many educators, the drive for accountability exemplifies the kind of top-down approach to educational change that undermines reflective practices in teaching and learning. Other educators see accountability as a necessary part of current efforts to reform schools. At the heart of the debate is the widespread recognition that, even if external authorities establish higher standards and provide inducements, many schools will still lack the organizational capacity to get their students to achieve at high levels (Newmann et al., 1997). As noted previously, many schools lack the necessary resources to respond to the needs of increasingly diverse student populations. Therefore, when schools are held accountable for ensuring that all students achieve high standards of learning, many complex issues are raised about the inequities that exist among schools.

Widespread concerns also exist about the use of penalties to enforce accountability when standards-based assessment systems are used. Accountability is an essential stimulus in achieving the scope of change needed in schools, but it is far more complex to implement than it is possible to define in typical state mandates. While schools owe their constituencies honest accounts of what they have and have not achieved, narrow visions of accountability can result in assessment being driven too exclusively by concerns for

Achieving Equity Through Standards and Assessments 19

reporting achievement data to external audiences. In the early stages of the current accountability in education, Wolf, LeMahieu, and Eresch (1992) emphasized the importance of internal accountability—encouraging students, teachers, and families to reflect on what is worth knowing, and ensuring that all students have the opportunity to develop essential knowledge. Internal accountability thus becomes as important as external mandates in addressing equity and excellence in the learning opportunities provided to students.