1

The Assessment Revolution

U.S. educators have been thrown into a score-boosting game they cannot win. More accurately, the score-boosting game cannot be won without doing educational damage to the children in our public schools.

—W. James Popham (2001, p. 12)

THE NEED FOR CHANGE

There is a growing recognition that, in spite of a recent rise in standardized test scores, students are not necessarily becoming better educated or more knowledgeable. The discrepancy between traditional **testing** results and the reality of what high school students know and can or cannot do provoked disenchantment with traditional methods of assessing student learning and has given impetus and momentum to the authentic **assessment** movement. While grades are here to stay in most districts, and state and national tests will continue to constitute a powerful external force in our school environments, what teachers and schools can do for the benefit of students is to counterbalance these standardized assessments with ongoing meaningful, authentic assessments of student understanding. An assessment focus that de-emphasizes traditional grades and demystifies the entire grading process is more in line with the manner in which the brain learns new information. Most advocates of authentic assessment do not want to totally eliminate traditional tests. What they seek is a good mix of assessment practices.

Cream of the Crop

Who is expected to learn? The central purpose for schooling has been reflected in this question. Historically, education has sorted and selected students much like separating cream from milk. While the "cream" was destined for higher education, the "milk" became the workforce required to power the industrial economy. Consequently, policies, practices, programs, and procedures were developed in which all students were taught the same way (the teacher lecturing to passive students), given the same amount of time, and tested with assessments based on the **bell-shaped curve.** The system manufactured a large common labor force with workers identified by their failure to achieve.

A system whose main purpose is sorting and selecting is at odds with the concept of educating all students. Outrage at the inequity of a stratified and stratifying education system was what propelled the legislation of the No Child Left Behind Act of 2001 (NCLB) and accounts for its overwhelming support from both political parties. The sad irony has been that the emphasis on testing and the practice of "teaching to the test" that the legislation has produced have undercut the values of the original intention. Brain-compatible, standards-based education introduces an orientation for ensuring that no child is left behind that is focused specifically on sustained achievement for all students— regardless of gender, race, ethnicity, or socioeconomic status. Given appropriate time, multiple and varied instructional strategies, and clear goal expectations, many of the students previously considered underachievers (the "milk") can now meet or exceed rigorous academic standards.

THE STANDARDS

The National Commission on Excellence in Education's now famous report *A Nation at Risk* sparked, or at least rekindled, the standards movement in America. Educational practices and policies had brought about what the report called "unilateral educational disarmament" (National Commission on Excellence in Education, 1983, p. 5). In an attempt to "re-arm" America's youth, a system driven by standards was sought. The Education Summit in 1989, called for and attended by then-president George H. W. Bush and the nation's governors, set the stage for the establishment of national standards. The National Council of Teachers of Mathematics (NCTM) soon developed a set of goals that were very well received. Other subject-area organizations followed NCTM's lead and developed national standards of their own. The National Education Standards and Improvement Council (NESIC) was created in 1994 as part of the Goal 2000 legislation to "oversee and certify" standards created by the states (Kendall & Marzano, 1996).

Soon however, the enthusiasm for standards waned. **Content standards** in some areas had sparked controversy, and the sheer size and complexity of the standards produced were overwhelming. In addition, by 1995 the NESIC had effectively ceased to function. The standards movement was down but not out. The re-authorization of the Elementary and Secondary Education Act (now

known as the Improving America's Schools Act [IASA]) in 1994 reinvigorated the standards dialogue by mandating the establishment of standards for schools receiving federal funds. By January 1998, 38 states had drafted academic standards in core subjects (English, math, science, and social studies) and 34 states used standards-based assessments of math and English. However, scholars engaged by the Thomas B. Fordham Foundation found that only one state had truly rigorous and clear standards in English, one in history, three in geography, three in math, and six states in science. A consistently rigorous level of standards in content area is a requirement if U.S. education is to step back from the precipice of risk. For educational reform based on standards to be successful, it must be consistent with brain-compatible instructional methodologies.

NCLB (2001), the most recent iteration of the Elementary and Secondary Education Act of 1965 (ESEA), introduced the issue of accountability. According to the U.S. Department of Education, accountability is a crucial step in addressing the achievement gaps that are perceived to exist. Under NCLB, every state is required to (1) set standards for grade-level achievement and (2) develop a system to measure the progress of all students and subgroups of students in meeting those state-determined, grade-level standards.

In effect, however, what NCLB has done is put an unreasonable emphasis on high-stakes test results, forcing teachers to forgo "meaningful and relevant" instructional methodologies in favor of a test-prep curriculum. Such high-stakes, statewide achievement tests do not measure the vast amount of curricula set forth by states and school districts. These tests tend to measure those things that are easy to measure, in an efficient and economical way. This means that the focus is on lower-order thinking skills, with a sprinkling of higher-order skills, such as writing a short essay (Popham, 2003). The reality of this situation is that schools and teachers, faced with ever-increasing demands to avoid the "failing school" label, logically focus on the curriculum content that is most likely to improve test scores, the unfortunate result being the narrowing of our nation's curriculum.

Writing and the ability to express oneself and one's environment are essential to literacy. Journal writing works with the processes of the brain and therefore stimulates understanding. The OPPORTUNITY FOR ASSESSMENT #1 is presented as a stimulus for introspection and cross-content knowledge integration. While creative writing has traditionally been a part of language arts instruction, creativity and writing are valuable in promoting student understanding of other content areas. When journals are used in multidisciplinary situations, they help build specialized vocabulary and a deeper understanding of underlying concepts. Students can keep a science fiction writing journal where the concepts experienced in the lab can be synthesized into an environment of the students' own making (a science fiction story). Or, students might engage in a historical diary project where they imagine themselves in the position of an historical figure, writing from that person's perspective.

The Creative Writing Journal is appropriate for students of every level as long as the performance criteria are appropriately adapted. Even emergent readers and writers can take part with the assistance of a "scribe" or a voice-activated word-processing program or tape recorder.

OPPORTUNITY FOR ASSESSMENT #1

Creative Writing Journal

Level Secondary

Content Standards

Language Arts: National Council of Teachers of English/International Reading Association (NCTE/IRA)

- ❑ Students adjust their use of spoken, written, and visual language (e.g., conventions, style, and vocabulary) to communicate effectively with a variety of audiences and for different purposes.
- ❑ Students employ a wide range of strategies as they write and use different writing process elements appropriately to communicate with different audiences for a variety of purposes.
- ❑ Students apply knowledge of language structure, language conventions (e.g., spelling and punctuation), media techniques, figurative language, and genre to create, critique, and discuss print and nonprint texts.

Multiple Intelligences and Learning Styles

❑ Visual/Spatial	❑ Bodily/Kinesthetic
❑ Logical/Mathematical	❑ Interpersonal/Social
☑ Verbal/Linguistic	☑ Intrapersonal/Introspective
❑ Musical/Rhythmic	❑ Naturalist

Performance Task

Over the course of a semester, students will maintain a personal creative writing journal in which they free-write for a half hour every day.

Assessment Technique

Portfolio creation; the journal itself serves as the portfolio. Only Novice (work is of poor quality) and Proficient (work is of satisfactory quality) are given because, in this instance, it is not what is being written that is being evaluated, but the fact that writing has taken place (see Figure 1.1).

Journal Assessment Rubric

Qualities Evaluated	Proficient Excellent and proficient work quality	Novice Beginning-level work quality
Effort	Journal entries for most days. Substantial material presented in each entry.	No journal entries for most days. Length of entries is insubstantial.
Individual Progress	Progress evident from first entry to last.	No progress evident from first entry to last.

Figure 1.1

FORMS OF ASSESSMENT

The educational system is now presented with the challenge of developing appropriate and meaningful ways to evaluate the extent to which students are meeting the standards. Time-honored methods, such as the standardized test, are being called into question. More progressive, brain-compatible methods of measurement are gaining favor in light of scientific information about ways in which the brain might be processing information. An example of how a **rubric** can be used to assess with a deeper view into a wide range of student performances can be seen in Figure 1.1. Such a rubric also provides a guide for assessment-driven instruction and a transparent guide for students that allows teachers to clearly communicate to learners the measures of success and allows students to monitor their own progress.

The Standardized Test

Horace Mann, the father of public education in the United States, was a firm believer in the concept of testing. As early as 1845, Mann was advocating the use of written tests comprising large numbers of questions with set answers as a way to evaluate student performance. Even then, early research in assessment and **evaluation** revealed problems with the objectivity of such evaluations. In further studies, it was discovered that when different teachers were given the same tests to grade, the scores varied widely and bore no relation to any uniform set of standards. In other early studies, papers that had received passing grades from evaluators the first time were graded as failures when given to the same evaluators a second time, with the reverse situation occurring with as much frequency (Hart, 1994).

The results brought attention to the need for greater objectivity in assessment. In fact, it was the perceived need for objectivity that first gave rise to the scientific testing movement early in the twentieth century. The direct result of the movement was the development of the standardized test, a mass-produced, multiple-choice test that could be administered to large numbers of persons with consistent results. To ensure the accuracy of the tests, statistical concepts and techniques were designed and developed to eliminate problems like subjectivity in scoring. By 1928, there were over a thousand standardized tests available in the United States, each with a statistically calculated measure of the test's validity (how well it measures what it is supposed to measure) and reliability (how consistently the scoring results will be over time and in different testing situations). The development of the standardized test led to the rise of the testing industry whose market expanded in the 1920s with the onset of official state testing. After World War II, two main trends contributed to the further expansion of the test industry: the decline of the one-room schoolhouse and accountability. By 1950, the one-room schoolhouse had evolved into a new model of instruction, one based on the factory design with its concept of mass production. The secret to making this factorylike schoolwork was to break the learning down into small skills and bits of knowledge that could be taught and learned sequentially as the students moved along the educational assembly line. Standardized tests fit nicely into this model of

instruction. With their multiple-choice formats, standardized tests made the perfect tool for measuring the mastery of subskills, or bits of content, for large numbers of students. With the advent of computerized scoring, standardized tests became even cheaper and easier to use.

The second trend that encouraged the growth of the test industry was the increasing concern over accountability. During the postwar years, baby boomers swelled school populations, and the amount of money that was spent on education increased precipitously. Unfortunately, student achievement did not keep pace with the educational investment. By the 1970s, Scholastic Aptitude Test (SAT) scores began falling while employers started complaining that high school graduates were neither able to read nor write. Angry taxpayers, not quite sure of where to place the blame for the failure, sought solutions by measuring and monitoring what went on in the schools. As a result, the standardized test became the assessment tool of choice for monitoring school quality.

By the 1980s, educators had begun to realize that the standardized test was fast becoming the driving force behind the curriculum. The consensus among many educators was that to move education away from its emphasis on memorization of isolated facts, education would need to move in a direction of increased emphasis on higher cognitive thinking skills. For such a change to take place, assessment methods needed to evolve and be redesigned into something other than the machine-scored multiple-choice test. With NCLB (2001) however, the politicized pendulum was demonstrating a return to those much-maligned standardized tests, to the chagrin of most educators.

Technically Speaking

Standardized tests usually fall into two categories: norm-referenced and criterion-referenced tests. Both test types rely heavily on the multiple-choice format. Norm-referenced standardized test scores give numbers that reflect achievement and performance of isolated skills at a particular moment in time. The norm-referenced testing is designed as a means of ascertaining an individual's performance in relation to the performance of other individuals on the same test. The scores are plotted along some form of the normal distribution curve. Such a test is designed to show how a given student or group of students rank when compared with other test takers of the same age and grade. With this method of evaluation, a certain percentage of the tested population must fail to establish the norm against which all the other students are measured. In a standardized test, the norm (or standard of performance) is determined by recording the scores of a large group, such as a sample of elementary school students. When subsequent students take the test, the norms for them will be those of the larger group (the group on which the test was "standardized"). Figure 1.2 shows a normal distribution curve. Test items are selected based on their ability to make distinctions among students. Items on which most students score either very high or very low are not retained for future tests because they do not discriminate among students. Consequently, some content standards may not be measured.
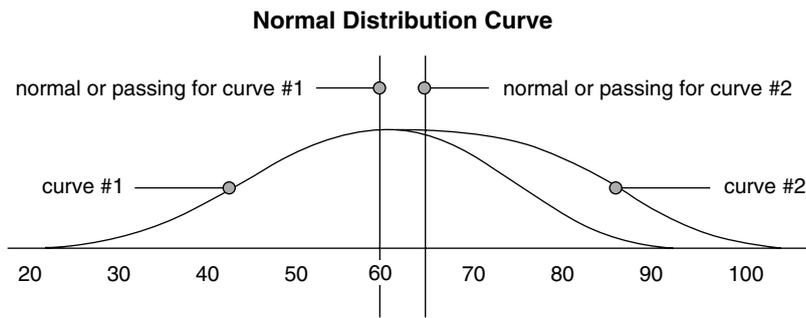
Normal Distribution Curve

Figure 1.2

A criterion-referenced test, on the other hand, is a way to measure an individual's performance against established criteria or expected standards (what a student should know at a given point in his or her education). The goal is to help all students attain at least the minimum level of mastery. Items chosen for inclusion on such tests are intended to reveal a student's strengths and weaknesses in terms of knowledge or skills. Competency tests and achievement tests are types of criterion-referenced tests. The criterion-referenced approach specifically provides for the following:

1. Test takers know what level of performance is expected of them in advance; therefore, students know what they must do to be successful on assignments and tests.

2. Teacher's subjective judgment is limited because the test content is not decided by a single individual.

3. Students are not fighting against each other for a place on the distribution curve; consequently, a noncompetitive environment results.

The problem with criterion-referenced tests from a brain-compatibility perspective is that the items chosen may be measuring only superficial knowledge, involving memorization and mechanics rather than higher-order thinking skills. In addition, because the evaluation occurs in an environment foreign to the subjects' context, is stressful, and occurs only once, it cannot give an accurate portrayal of an individual's knowledge.

Critics point out that standardized tests corrupt the very process they seek to improve. America's reliance upon tests, they contend, has damaged teaching and learning by putting too much value on memorization of isolated bits of information at the expense of higher thinking skills. Teachers often feel the pressure to focus more on what can be easily tested than on what is important for students to learn. Students are thereby conditioned to be passive learners who can only recognize information rather than construct their own answers and solutions.

In addition, standardized testing neglects the vital aspect of emotion in assessment. (For a discussion of the role emotion plays in the learning process, see the "Emotion in Context" section of Chapter 4.) If a student typically learns in a classroom and then is tested in a media center or an auditorium, that student is more likely to underperform. Similarly, if a student learns in a particular emotional state, he or she will most readily recall that learning when in that same state. It is the job of the assessor to match the memory mechanism at assessment time to that which occurred during learning, or the student may not be as successful as he or she might be in demonstrating the knowledge learned. The best way to do such an evaluation is by providing opportunities for assessment that occur as part of the learning process.

Observation and conferencing strategies work very well for tasks that have some kinesthetic or performance element to them where progress toward a goal can be monitored in addition to assessing the final product. OPPORTUNITY FOR ASSESSMENT #2 has students acting out a kind of interplanetary play. Planets in Motion offers a performance task appropriate to the elementary level. (Pre–K students may also be able to learn from and enjoy this activity.) In addition, middle-level students can perform the task with a higher degree of accuracy and "reality" by plotting the respective planets' orbits and scaling them down to fit on the playground. They may also be asked to reflect upon the role size plays in the gravitational pull of an object and why planets travel in elliptical orbits. Secondary-level students may act out more complex processes or re-create famous experiments "live" in front of an audience. These performances can be videotaped and used by students to critique their individual performances, or the tapes can stand as example for future classes of what excellence looks like. Activities in which students play a particular role in a complex process must allow time for students to see their role as part of a bigger picture, otherwise learning may be fragmented. Witnessing their performance on videotape or seeing their place on a diagram of their creation will help them bring their experiences into the larger context.

Observation Rubric

Focus Items	Observations
Concepts: planet placement in the solar system	Jimmy has difficulty locating Earth's position relative to the sun. Beth (as the sun) explained to Jimmy why he had to move around her.
Techniques: verbal interaction of students with teacher guidance	Some students are having difficulty with each planet's orbit and rotation speed. Design an activity that will help focus them on this concept.
Problem solving/reasoning	Some of the students are able to explain the logic of the solar system to the others. We need to go over the basic planet relationships in class again.
Communication and collaboration	Beth and Tom are very helpful with the other students, always trying to help them see where they might fit into the orbit design. Jimmy, as Earth, is set on being the center even though he knows his science facts.

Figure 1.3

OPPORTUNITY FOR ASSESSMENT #2

Planets in Motion

Level Elementary

Content Standards

Science

Content Standard D: Earth and Space Science

- ❑ Structure and properties of the Earth system
- ❑ Earth's history and cycles
- ❑ Earth in the solar system

Language Arts

- ❑ Adjusts use of spoken, written, and visual language (e.g., conventions, style, and vocabulary) to communicate effectively with a variety of audiences and for different purposes

Physical Education

- ❑ Demonstrates competency in many movement forms and proficiency in a few movement forms
- ❑ Applies movement concepts and principles to the learning and development of motor skills

Multiple Intelligences and Learning Styles

☑ Visual/Spatial	❑ Bodily/Kinesthetic
❑ Logical/Mathematical	❑ Interpersonal/Social
☑ Verbal/Linguistic	❑ Intrapersonal/Introspective
❑ Musical/Rhythmic	❑ Naturalist

Performance Task

In a large area such as a playground, students will take turns placing themselves in different locations as they act out the rotation of the planets and their orbits. Once back in the classroom, students will write a paragraph about the planetary relationships.

Assessment Technique

Performance assessment (informal observation). Observation is an important means of assessment integration. With observation, the teacher is able to monitor the learning process as it occurs within the group (see Figure 1.3).

Student Samples (see Figure 1.5)

Paragraph Assessment Rubric

CRITERIA EVALUATED	NOVICE BEGINNING 1 NOT YET	BASIC DEVELOPING 2 YES BUT	PROFICIENT ACCOMPLISHED 3 YES	ADVANCED EXEMPLARY 4 YES PLUS
COMPREHENSION				
Level of Comprehension	Weak demonstration of comprehension; vague and confusing	Inconsistent demonstration of comprehension; superficial	Solid and consistent demonstration of comprehension	Masterful demonstration of profound and sophisticated comprehension
THOUGHT ORGANIZATION				
Ideas Clearly Focused and Well Supported	Few ideas are focused or supported	Ideas are focused and supported in places, but lack consistency	Ideas are competent, focused, and well supported	Ideas are powerfully supported, consistently focused, perceptive, and comprehensive
DEMONSTRATION OF EFFECTIVE WRITING				
Organization	Writing is aimless and disorganized	Writing demonstrates simplistic organization	Writing demonstrates logical and competent organization	Writing demonstrates mature and sophisticated organization
Word Choice	Repetitive, dull, or incorrect word usage	Routine, simple word choices	Some word choices are sophisticated, while others are routine	Word choices are vivid and sophisticated yet natural
Sentence Fluency	Numerous awkward, run-on, and/or fragmented sentences, making reading difficult	Sentence structure is uneven	Writing demonstrates good control of sentence structure	Writing demonstrates powerful and sophisticated control of sentence structure
EVIDENCE OF CRITICAL THINKING				
Clarity of Thought Analysis	Thoughts are vague, nonspecific, and lack evidence of organization	Ideas are logical but not well analyzed	Several ideas are logical and have been well analyzed	Intuitive analysis, with clear and precise thought process throughout

Figure 1.4

Student Samples

Advanced	Proficient
For over science period, we went to the playground and pretended to be the planets in our solar system. Mr. Blank told us to try and rotate in our own orbits as if we were the real planets moving around the sun.	Today we played that we were the planets turning around the sun. We were turning in paths called orbits.
Basic	**Novice**
We went to the playground today and made beleave we were the planits. I wanted to be the second plant venis.	We play planet in yard. I play mercry. I go to sun.

Figure 1.5

ALTERNATIVE ASSESSMENTS

Alternative assessments are any and all assessments other than the standardized test–type assessments. A broad definition of alternative assessment includes any type of assessment in which a student creates a response to a question rather than choosing a response from a given list (as with multiple-choice, true/false, or matching). Some of the different alternative assessments include short-answer questions, essays, products, performances, oral presentations, demonstrations, exhibitions, and portfolios.

Standards found in NCTM's *Principles and Standards for School Mathematics* (2000) and the National Research Council's *National Science Education Standards* (1995) present a vision of assessment that is highly brain compatible in that it is ongoing and carried out in multiple ways. By listening to, observing, and talking with students, by asking students questions to help reveal their reasoning, by examining students' individual or group written and/or problem-solving work, teachers are able to develop a more accurate and valid picture of what students know and can do.

When conceived of and used in such a nonthreatening, brain-compatible manner, assessment provides teachers with the best way to gain valid insights into their students' thinking and reasoning abilities. Consequently, assessment becomes a powerful tool to help teachers monitor the effectiveness of their own teaching, judge the utility of the learning tasks, and consider when and where to go next in instruction. The focus is on high but achievable targets for students to meet rather than comparing student performance to the performance of other students. When information is shared with the students regarding precisely what they are expected to know and be able to do, students are able to meet or even exceed the standards.

Informal and Formative Assessments

Formative assessments consist of information, gathered by teachers mainly in their day-to-day classroom encounters with students, that registers students' internal processing of information, the development of student understanding,

student-to-student interaction, and the discussion and revision of ideas. This process can take a variety of formats and is used to inform the teacher's instruction and pinpoint needs of the group and of individual students to support their learning progress. Assessment should be embedded in instruction, or in other words, integrated within the instruction. To evaluate the intangible processes taking place in the classroom, especially during group activities, teachers gain important information from observations. Observation is particularly important as a means of bringing about such integration of instruction and assessment. Through group as well as individual student observations, teachers are able to develop a sophisticated and complex picture of student learning. By using an observation time frame of 10 to 15 minutes, teachers can acquire the optimal perspective on group processes regarding topic content, social dynamics, or needed help. The Observation Rubric (see Figure 1.3) provides a sample of how teachers might officially record their assessment observations in an orderly and relatively objective manner. It is a good method for the organization of observations and helps to connect students to the learning content.

Another informal assessment method is to conduct active meetings with students either individually or in small groups. Such meetings, or conferences, can be brief (3 to 5 minutes) or longer if needed. It is often beneficial to give a particular student or group a specific assignment, such as exploring a problem, asking a peer for an explanation, or researching a similar topic, in order to provide a focus for student work. In this way, the next conference or observation held with that particular group or individual will provide the teacher with a good indication of the progress made by the group or individual. Focused, informal assessments help teachers keep up with students' varied progress rates and needs, which is especially important with more extensive projects.

Authentic Assessment

One of the major forms of alternative assessment is the authentic assessment. An assessment is considered to be "authentic" when it involves students in tasks that are worthwhile, significant, and meaningful (tasks that take into account that the search for meaning is innate in all humans); occurs over time (not just one day); is open-ended; and allows for students to demonstrate competence in a variety of ways. Such assessments involve higher-order thinking skills along with the use of a broad range of knowledge. In addition, authentic assessment demonstrates to the student exactly what it means to do excellent work by making explicit the standards by which that work will be evaluated. In this sense, authentic assessments are standard-setting assessment tools rather than standardized assessment tools.

Authentic assessments are brain compatible. When correctly designed, they emphasize learning and thinking, especially those higher-order thinking skills involved in problem solving. Authentic assessments comprise meaningful tasks that reflect real-life, interdisciplinary challenges; they present students with complex, ambiguous, open-ended problems and tasks that integrate their knowledge and skills. Such assessments usually culminate in student products

Standardized Tests vs. Brain-Compatible Assessment

Standardized Testing	Brain-Compatible Assessment
• results based on a mythical standard or norm, which requires that a certain percentage of children fail	• establishes an environment where each child has the opportunity to succeed
• pressures teachers to narrow their curriculum so that they can specifically concentrate on the test material	• allows teachers to develop meaningful curricula and assess within the context of that program
• emphasizes a single instance assessment, which has no relation to the learning taking place in the classroom	• assessment is ongoing throughout the unit of study, and provides an accurate picture of student achievement
• focuses on errors and mistakes rather than what has been accomplished	• puts the emphasis on student strengths rather than weaknesses
• focuses too much importance on single sets of data (i.e., test scores) in making educational decisions	• provides multiple sources of evaluation that give an in-depth view of student progress
• treats all students in a uniform way	• treats each student as a unique human being
• discriminates against some students because of cultural background and learning style	• provides the opportunity to eliminate cultural bias and gives everyone an equal chance to succeed
• regards instruction and assessment as separate activities	• regards instruction and assessment as being a single, integrated activity
• answers are final, there is no opportunity for reflection or revision	• engages the student in a continual process of self-reflection, learning, and feedback, as well as revision
• focuses on the "right" answer without regard for understanding	• deals with comprehension and the learning process as much as the final product
• inexpensive and easy to administer and grade	• more difficult to achieve consistent, objective scoring results
• often provides results that can be simplified to a single numerical score	• data cannot easily be simplified as a single number
• easy to compare and contrast different populations of students	• difficult to compare different student populations

Figure 1.6

or performances that recognize and value each student's multiple abilities, varied learning styles, and individual background. Calculators and dictionaries are necessary tools in the real world. If the job of educators is to prepare students to function in the real world, then students should be taught when and how to use real-world tools. Therefore, calculators, dictionaries, relevant textbooks, and other materials should be made available to students during an evaluation if the evaluation is genuinely authentic.

To be brain compatible, authentic assessments must have the following qualities:

Structure

- Involve an audience, are more public than traditional forms of assessment
- Do not rely on arbitrary or unrealistic time constraints
- Contain questions or tasks that are known beforehand and are not "secret"
- Encompass multiple opportunities for demonstration of growth (i.e., portfolios) rather than one-time, stressful experiences
- Include some sort of collaboration with peers
- Allow for a significant degree of student choice

Intellectual Design Features

- Direct students toward more sophisticated uses of knowledge and skills (i.e., critical thinking skills)
- Integrate tasks and their outcomes
- Assess thinking processes rather than bits and pieces of isolated information
- May involve somewhat ambiguous or "messy" tasks and/or problems to be solved
- Utilize the student's own research or use of knowledge
- Present a challenge that emphasizes depth of knowledge and understanding
- Stimulate and educate so that students can learn from the assessment process

Grading and Scoring Standards

- Are based on clearly articulated criteria and **performance standards** rather than a curve or norm
- Use performance indicators, which allow students to know ahead of time what excellence looks like
- Make metacognitive activities such as self-assessment and self-reflection part of the total assessment process
- Use a multifaceted scoring system rather than a single numerical grade
- Demonstrate equity
- Identify hidden strengths rather than weaknesses
- De-emphasize competitive comparisons between students
- Allow for different learning styles, abilities, and interests

A STARTING POINT

The results of standardized norm-referenced tests have become the virtual standards by which American public education is judged. Standardized testing has its place, but it is ultimately a poor judge of a school's quality of instruction. Relevant and meaningful standards in conjunction with authentic assessments would provide a real means of evaluation. Any attempt to change the status quo, however, would first require that the validity of a new assessment strategy be quantified. A starting point for such would be the comparison of the performance of the same group of students on both the norm-referenced test and the replacement measure. Standards-based or criterion-referenced assessments can be cross-referenced to standardized norm-referenced tests by conducting a concurrent validity study (Burger & Burger, 1993). In such a study, students would complete both a standardized norm-referenced test and a criterion- or standards-based test measuring a similar domain, reading for example. Statistical analysis could then provide both correlated data and the relationship of performance standards to percentile scores. The linkage between the two assessments would assure that the new assessments are rigorous and that the performance standards are worthy. There is no one best assessment method; hence the flaw in NCLB. The choice of using multiple-choice tests, performances, projects, exhibitions, or portfolios depends on what is being assessed, the purpose for the assessment, and how the assessment results will be used. A combination and variety of assessment formats is what is needed for a thorough and complete picture of student progress and growth.

A project that takes a number of weeks to complete will, in theory, allow for greater student thought about the process and what they want their product to look like. Portfolio-type performance tasks, such as the one presented in OPPORTUNITY FOR ASSESSMENT #3, Campaign Scrapbook, allows students to choose the way in which they will depict an idea or concept. When left to choose, students will often pick depictions and activities that are best suited to their most pronounced learning styles. The end result of a campaign scrapbook will look different for each student, but the criteria upon which it is evaluated will remain the same. The organizational structure, for instance, may vary widely, but some organizational planning should be evident in the product. The scrapbook concept can be adapted to suit additional content areas and levels. Elementary students can create a scrapbook of food labels and remark upon the use of color, wording, and logos to introduce or reinforce the concept of the symbolic use of language. Secondary students might keep the campaign element of the scrapbook but include editorials along with campaign ads to compare and contrast the approaches, the purpose, and the relative effectiveness of each. Student can then reflect upon the concept of rhetoric and debate. An extension of the task at all levels can include students creating their own advertisements (or product labels).

OPPORTUNITY FOR ASSESSMENT #3

Campaign Scrapbook

Level Middle

Content Standards

Social Studies: National Council for the Social Studies (NCSS)

Individuals, Groups, and Institutions

- ❑ Social studies programs should include experiences that provide for the study of interactions among individuals, groups, and institutions.
- ❑ In schools, this theme typically appears in units and courses dealing with sociology, anthropology, psychology, political science, and history.

Power, Authority, and Governance

- ❑ Social studies programs should include experiences that provide for the study of how people create and change structures of power, authority, and governance.
- ❑ In schools, this theme typically appears in units and courses dealing with government, politics, political science, history, law, and other social sciences.

Language Arts

- ❑ Students read a wide range of print and nonprint texts to build an understanding of texts, of themselves, and of the cultures of the United States and the world; to acquire new information; to respond to the needs and demands of society and the workplace; and for personal fulfillment. Among these texts are fiction and nonfiction, classic and contemporary works.
- ❑ Students adjust their use of spoken, written, and visual language (e.g., conventions, style, and vocabulary) to communicate effectively with a variety of audiences and for different purposes.
- ❑ Students employ a wide range of strategies as they write and use different writing process elements appropriately to communicate with different audiences for a variety of purposes.
- ❑ Students apply knowledge of language structure, language conventions (e.g., spelling and punctuation), media techniques, figurative language, and genre to create, critique, and discuss print and nonprint texts.

Multiple Intelligences and Learning Styles

☑ Visual/Spatial	❑ Bodily/Kinesthetic
❑ Logical/Mathematical	❑ Interpersonal/Social
☑ Verbal/Linguistic	❑ Intrapersonal/Introspective
❑ Musical/Rhythmic	❑ Naturalist

Performance Task

Students will maintain a scrapbook of campaign advertisements from a variety of different candidates and chronicle alongside the ads their observations and reflections on those advertisements. Students can compare and discuss collected artifacts. This discussion can support individual learning and scaffold English language learners. In the case of the latter, some students may participate primarily at the level of collecting based on recognizable photos and names in headlines. This is an entry point to learning and group discussion, particularly with peer translating, where possible, to support their differentiated learning.

Assessment Technique

Portfolio Assessment is perhaps the best method to show a student's processing and reflection (see Figure 1.7). A rubric, such as the one in Figure 1.7, is a valuable portfolio assessment tool because it shows the student exactly where the artifact is strong and where it needs to be improved.

Campaign Scrapbook Evaluation Rubric

CRITERIA EVALUATED	NOVICE BEGINNING **1 NOT YET**	BASIC DEVELOPING **2 YES BUT**	PROFICIENT ACCOMPLISHED **3 YES**	ADVANCED EXEMPLARY **4 YES PLUS**
Organization	Poor organization with parts difficult to find or missing completely	Good organization, but not all parts included or not readily accessible	Well organized	Sophisticated organization; all parts included and readily accessible
Communication of Ideas	Ideas presented in a confusing manner	Some ideas communicated well, but others are confusing	All ideas clearly communicated	All ideas clearly communicated in a sophisticated and original manner
Comprehension (Insights and Connections)	Demonstrates little or no understanding	Some understanding demonstrated, but connections are limited	Clear connections demonstrated	Exceptional insight demonstrated by sophisticated connections
Variety	Little or no variety in ad choice	Some good ads included, but variety is limited	Good variety in choice of ads	Numerous high quality ads chosen from a wide variety of sources
Presentation	Insufficient effort is obvious; needs more work	Satisfactory presentation, but nothing to make it stand out	Attractive presentation catches audience attention	Creative and unique presentation; stands out from the rest
Reflection	Little or no evidence of reflective thought processes	Some insight evident, but inadequate effort	Evidence of insight and constructive thought	Mature, thorough, realistic, and constructive

Figure 1.7