

## **ASSESSMENT: THE KEY TO STANDARDS IMPLEMENTATION**

Standards-based education promotes the idea that content, instruction, and assessment must be aligned to provide students with a coherent, cohesive experience in school. That experience should prepare students to enter the world of work or proceed to postsecondary education.

The first steps in developing standards-based education are determining appropriate standards; then articulating them to students, parents, potential employers and all member of the entire educational community. The instruction and assessment are then based upon these agreed upon standards.

### **STANDARDS-BASED ASSESSMENT SYSTEM**

Standards-based assessment can be a powerful tool for teachers and students for the following reasons:

- The standards set clear targets and expectations for students, teachers, and parents.
- Students improve their ability to monitor and evaluate their own work and assess their progress as they work with standards-based sources of evidence.
- Products and performances provide an ongoing, cumulative data source for measuring progress over time, within and across units.

California's assessment system is built upon the assessment of high quality standards. Career and Technical Education is a key component of the state's vision for student assessment. Other components of the system include:

- Statewide standards and assessments
  - Assessment in Career Education (ACE) – The ACE program, patterned after the well-established Golden State Examination program is designed to recognize students who demonstrate outstanding achievement on rigorous examinations in selected career-technical areas. All examinations are based on career education standards. The examinations assess students’ knowledge of the subjects and their application of that knowledge. Examinations include multiple-choice and written-response questions.
  - STAR Accountability Program – STAR is the Standardized Testing and Reporting program. The STAR program has three components: the Stanford Achievement Test, Ninth Edition, Form T (Stanford 9), the STAR augmented tests for English/language arts and mathematics, and the Spanish Assessment of Basic Education (SABE/2)
  - Stanford 9 – The Stanford 9 is a nationally normed multiple-choice achievement test. School districts are required to administer the Stanford 9 to all students in grades 2 through 11. A combination of multiple-choice and open-ended subtests helps obtain a more complete picture of both the breadth and depth of students’ educational achievement.
  - High School Exit Exam – State law (Senate Bill 2), passed in 1999, authorized the development of the California High School Exit Examination (CAHSEE) that students in California public schools will have to pass to receive a high school diploma, beginning with the graduating class of 2004. The purpose of the CAHSEE is to improve student achievement in high school. It also is to help ensure that students who graduate from high school can demonstrate competency in state content standards for reading, writing,

and mathematics. The Language Arts and Mathematics Blueprints for the CAHSEE are located under TAB 10.

- Golden State Examinations (GSE) – The GSE Program was established in 1983 to offer rigorous examinations in key academic subjects to students in grades 7 – 12 and to recognize students who demonstrate outstanding achievement on each examination. The examinations assess students’ knowledge of the subjects and their application of that knowledge. Examinations include multiple-choice and written-response questions.
- Local standards and assessments
  - Expected Schoolwide Learning Results (ESLRs) – ESLRs are the school’s statements of what students should know and be able to do upon graduation. It is the school’s vision for all students. They are interdisciplinary and all disciplines can support student learning of the ESLRs in their courses.
  - District and Site Assessments – Tools and processes used by a district or site to gather data about student results and the implementation of programs and practices. The extent of these assessments will vary district by district.
- Various types of assessments
  - On-demand assessments - *On-demand assessments* literally demand that students show their knowledge on a test at a predetermined time and place, under uniform conditions. On-demand assessments can include multiple choice, true/false, short-written responses, even short laboratory experiments, but by their nature are designed to be relatively short. Selected-response questions (multiple choice and true/false) are usually narrow enough in focus to address discrete pieces of knowledge that address one

standard. Yet, a collection of such questions allows for broad coverage of many standards. One obvious benefit of selected-response questions is that the scoring of student responses is both quick and inexpensive. On-demand assessments can include the standardized, norm-reference tests that districts purchase and use to assess students yearly. They can also include district and teacher constructed test, but they provide information about the ability of a student to answer questions on a certain day.

- Cumulative assessments - *Cumulative assessments* usually assess a student's acquisition of standards to a greater depth. Because of their scope, these long-term assessments offer a more thorough picture of the student's ability to apply knowledge. Cumulative assessments usually require more time on the part of the students and are not as easily scored. Examples include portfolios, projects, exhibitions, and lengthy written scenarios. These cumulative performance assessments require an investment of instructional time to prepare students both in content knowledge and skills they need and to acquaint students with the form of the assessment. Consequently, the line between instruction and assessment is frequently blurred. The instruction that teaches students to reflect on they own progress might easily become part of the assessment itself.
- Performance assessments - *Performance assessments* require that a student actually perform the assessment rather than choose a correct response from several options offered. If, for example, an assessment hopes to test a student's ability to write a memo in regards to a business issue, the assessment might have the student actually write the memo. This would be a performance assessment. Performance assessment might be a part of either on-demand or cumulative assessments.

## STANDARDS-BASED ASSESSMENT

### On Demand Assessment

- Limits students to a certain time and place and uniform conditions
- Utilize questions that focus on discrete pieces of knowledge
- May cover a breadth of standards by asking many questions
  - Assessments in Career Education (ACE)
  - SAT 9
  - CAHSEE
  - Selected Response, True/False, Multiple Choice, . . .
  - Short Lab Experience

### Cumulative Assessment

- Allows for flexibility in student responses
- Captures student progress overtime
- Assesses to a greater depth
- Requires more time
  - C-TAP Model
    - Portfolios
    - Projects
    - Lengthy written scenarios

There are dramatic differences between the traditional approach and the standards-based approach to designing assessment and instruction. These pronounced differences are apparent in the chart on the following page that provides a comparison of traditional and standards-based assessment.

## Traditional Assessment Vs. Standards-based Assessment

### Traditional

- The assessment is secret
- Traditional tests are associated with national norms
- Traditional assessments seek to discriminate among different students
- Traditional assessments are overwhelmingly multiple choice tests
- Traditional assessments are typically limited in time

### Standards-based

- Standards-based assessments are open, not secret
- Standards-based assessments are designed so that a large number of students can achieve proficiency
- Standards-based assessments involve a demonstration of proficiency, not a guess on a multiple choice test
- Performance assessments recognize the fact that there is not a single “right” answer on a number of test items
- Standards-based performance assessments force educators to come to grips with this central question: “What do we expect of our students?”

\* *Making Standards Work*, Douglas B. Reeves, Center for Performance Assessment, 1996-1998.

Business Education teachers who are aligning their curriculum with the Challenge Standards (Tab 6) are ready to match the selected standards to the most appropriate form of assessment. Richard Stiggins of the Assessment Training Institute, Inc., in Portland, Oregon, has considerable expertise in matching achievement targets with assessment methods. The following information relies on his work.

Stiggins considers four different types of assessment:

- selected response (including multiple choice, true and false, matching, and fill in);
- essay;
- performance assessments; and
- personal communications.

He indicates that each type of assessment lends itself to particular targets. For example, selected response is an appropriate choice for assessing knowledge mastery. The chart on the following page provides more examples.

**THE VARIOUS POSSIBLE LINKS BETWEEN ACHIEVEMENT  
TARGETS AND ASSESSMENT METHODS**

TARGET TO BE ASSESSED	ASSESSMENT METHOD			
	SELECTED RESPONSE	ESSAY	PERFORMANCE ASSESSMENT	PERSONAL COMMUNICATION
KNOWLEDGE MASTERY	Multiple choice, true/false, matching and fill-in can sample mastery of elements of knowledge	Essay exercises can tap understanding of relationships among elements of knowledge	Not a good choice for this target—three other options preferred	Can ask questions, evaluate answers and infer mastery—but a time consuming option
REASONING PROFICIENCY	Can assess understanding of basic patterns of reasoning	Written descriptions of complex problem solutions can provide a window into reasoning proficiency	Can watch students solve some problems and infer about reasoning proficiency	Can ask students to “think aloud” or can ask follow-up questions to probe reasoning
SKILLS	Can assess mastery of the knowledge prerequisites to skillful performance, but cannot tap the skill itself—not a good choice for this target		Can observe and evaluate skills as they are being performed	Strong match when skill is oral communication proficiency; also can assess mastery of knowledge prerequisite to skillful performance
ABILITY TO CREATE PRODUCTS	Can assess mastery of knowledge prerequisite to the ability to create quality products, but cannot assess the quality of products themselves—not a good choice		A strong match; we can assess: (a) proficiency in carrying out steps in product development, and (b) attributes of the product itself	Can probe procedural knowledge and knowledge of attributes of quality products—but not product quality

From Stiggins, Richard J., *Student-Involved Classroom Assessment*, 3<sup>d</sup> ed. (Columbus, Ohio: Merrill, 2001)

The alignment of standards to assessment includes the process of deciding which assessment to use when. Although certain types of assessment may come and go in terms of popularity, some forms of assessment have strengths that better assess particular standards due to the nature of the standards.

A local standards-based assessment system will undoubtedly make use of all the various forms of assessment in such a way that the strengths of a certain kind of assessment can be tapped. Some forms of assessment require a great deal of instructional support and should be chosen only for those standards of relative importance.

An Assessment Task Selector is provided on the following page to assist teachers in deciding on an appropriate assessment model. The Assessment Task Selector also encourages teachers to think about a wide variety of assessments to use in the classroom. As teachers make decisions about what form of assessment to use when, they should consider:

- Does the standard lend itself to a particular assessment due to the nature of the standard?
- Does the time required for the assessment match the relative importance of the standard being assessed?

When determining the appropriateness of assessments, teachers should evaluate the assessments they are currently using and ask themselves whether these assessments fit the criteria for standards-based assessment. An Assessment Critique Sheet has also been provided in the following pages to assist in this process.

## ASSESSMENT TASK(S) SELECTOR

Developed by the SBE Design Team, Northern Colorado BOCES

Standard(s) Benchmark(s):

**INSTRUCTIONS:** After analyzing the standards and benchmarks to identify the specific knowledge and skills required, you are ready to select one or more assessment types and to design and write prompts (student instructions) for one or more assessment tasks. First, go through the lists below (with a creative and open mind) and check *all* of the types that *might* be used in this case. Then go back and circle the type or types that will work best. Consider using tasks that mirror real-world activities by adults when they employ this knowledge and these skills. Be sure that the task(s) you design will provide sufficient, valid evidence regarding the specific knowledge and skills you outlined previously. Consider also such aspects as feasibility, instructional value, and potential bias.

- |  |  |
|--|--|
| <input type="checkbox"/> Multiple Choice                           | <input type="checkbox"/> Documented Report                       |
| <input type="checkbox"/> True/False                                | <input type="checkbox"/> Poster; Visual Aid                      |
| <input type="checkbox"/> Multiple Choice with Supporting Rationale | <input type="checkbox"/> Slide Show                              |
| <input type="checkbox"/> True/False with Supporting Rationale      | <input type="checkbox"/> Audio Tape                              |
| <input type="checkbox"/> Matching                                  | <input type="checkbox"/> Interview                               |
| <input type="checkbox"/> Fill-In-the-Blank                         | <input type="checkbox"/> Discussion; Panel; Debate               |
| <input type="checkbox"/> Short Answer                              | <input type="checkbox"/> Simulation; Role Play                   |
| <input type="checkbox"/> Outline; Web                              | <input type="checkbox"/> Video Tape; Documentary                 |
| <input type="checkbox"/> Brainstorm List                           | <input type="checkbox"/> Multi-Media Presentation                |
| <input type="checkbox"/> Sorting of Manipulatives                  | <input type="checkbox"/> Student-Generated Problem with Solution |
| <input type="checkbox"/> Notes; Summary                            | <input type="checkbox"/> Story; Poem; Dialogue; Script           |
| <input type="checkbox"/> Time Line                                 | <input type="checkbox"/> Dramatization; Puppet show              |
| <input type="checkbox"/> Journal; Learning Log                     | <input type="checkbox"/> Choreographed Dance                     |
| <input type="checkbox"/> Album                                     | <input type="checkbox"/> Original Song; Music Composition        |
| <input type="checkbox"/> Map                                       | <input type="checkbox"/> Scale Model; Diorama; Exhibit           |
| <input type="checkbox"/> Graph                                     | <input type="checkbox"/> Three-Dimensional Construction          |
| <input type="checkbox"/> Flow Chart                                | <input type="checkbox"/> Interactive Computer Program            |
| <input type="checkbox"/> Illustration; Diagram; Mural              | <input type="checkbox"/> Investigation; Experiment               |
| <input type="checkbox"/> Table                                     | <input type="checkbox"/> Brochure; Pamphlet                      |
| <input type="checkbox"/> Graphic Organizer                         | <input type="checkbox"/> Commercial; Ad Campaign                 |
| <input type="checkbox"/> Letter; Editorial                         | <input type="checkbox"/> Questionnaire; Survey                   |
| <input type="checkbox"/> News Story                                | <input type="checkbox"/> Budget; Cost/Benefit Analysis           |
| <input type="checkbox"/> Filmstrip                                 | <input type="checkbox"/> Performance                             |
| <input type="checkbox"/> Essay                                     | <input type="checkbox"/> Portfolio                               |
| <input type="checkbox"/> Recipe; Set of Instructions               | <input type="checkbox"/> Other _____                             |
| <input type="checkbox"/> Technical Explanation                     | <input type="checkbox"/> Other _____                             |
| <input type="checkbox"/> Oral Presentation; Speech                 | <input type="checkbox"/> Other _____                             |
| <input type="checkbox"/> Demonstration                             | <input type="checkbox"/> Other _____                             |
| <input type="checkbox"/> Children's Book                           |  |
| <input type="checkbox"/> Instructional Materials                   |  |

## ASSESSMENT CRITIQUE SHEET

Developed by the SBE Design Team, Northern Colorado BOCES

Assessment Title: \_\_\_\_\_ Evaluator: \_\_\_\_\_

Instructions: Study the assessment carefully. Then to the left of each statement, circle **SA** if you **Strongly Agree** with that statement in relation to this assessment, **A** if you **Agree**, **D** if you **Disagree**, or **SD** if you **Strongly Disagree**.

- SA A D SD 1. The assessment is a valid measure of the standard. It accurately and comprehensively measures the knowledge and/or skills described in the standard. Where possible, it avoids the need for students to use knowledge or skills not related to the standard. Each student's performance level reflects his/her personal abilities.
- SA A D SD 2. The instructions are clear. The print, diagrams, and layout are clean and easy to follow. Students know exactly what they are asked to do and what criteria will be used to judge their work. Test administrators know the time, materials, and resources to give students and what kinds of help are appropriate and inappropriate.
- SA A D SD 3. The assessment is meaningful. Students would find the assessment relevant and be motivated to do their best work. The assessment provides an interesting challenge and would bring a sense of accomplishment if done well. It encourages and allows all students to do their best work. The task and context are the same as or an authentic simulation of the work done by adults.
- SA A D SD 4. The assessment is instructionally useful and worth teaching to. Students learn during and after the assessment and are able to improve their performance. The assessment gives the teacher useful information which guides good instruction.
- SA A D SD 5. The scoring system accurately reflects the essential characteristics of a good performance. Examples of excellent student work would earn high scores and examples of poor student work would earn low scores. Performance levels are clear, developmentally appropriate, and publicly credible. Distinctions described in the scoring guide are real, not arbitrary. Appropriate forms for scoring and reporting are included. Trained scorers would likely give the same rating to a given student performance.
- SA A D SD 6. The assessment avoids bias. It is unlikely that some students score higher than others due to differences in gender or ethnicity.
- SA A D SD 7. The assessment yields clear results understood by students, parents, and others.
- SA A D SD 8. The assessment is feasible. A realistic amount of time and resources is invested in taking and scoring the assessment. The results are worth the effort.
- SA A D SD 9. Appropriate accommodations of the assessment for special subpopulations of students are delineated. It is clear what changes in the assessment instructions and conditions can be made to facilitate student performance for some individuals without harming validity.
- SA A D SD 10. The assessment includes guidance for the teacher regarding the prior learning students should have. Knowledge and skill components of the assessment have been analyzed for the teacher, and suggestions have been given regarding effective instructional content and strategies to prepare students for this assessment.

From the of the *Career-Technical Assessment Program Teacher Guidebook*, WestEd Publisher:

Two of the assessment systems are Assessments in Career Education (ACE) and the Career-Technical Assessment Program (C-TAP). ACE is a state-level, voluntary testing program for high school students in California. It is being developed and administered by the California Department of Education (CDE), with the help of WestEd and the Sacramento County Office of Education (SCOE). Each ACE examination includes a combination of multiple-choice and written-response questions and is designed to assess the breadth of students' standards-based knowledge in a specific career area. In addition to measuring student achievement, each ACE examination can serve as one indicator of students' preparedness for the workplace and continued education and training.

C-TAP is another means of ensuring that high school students achieve mastery of standards-based knowledge and skills in a specific career area. C-TAP is an assessment system developed by WestEd for implementation at the local level (i.e., by individual teachers, schools, and school districts). The objectives of C-TAP are two-fold: 1) to help California students learn important career-technical knowledge and skills, and 2) to assess student readiness for entry-level jobs and post-secondary educational training.

### **Key Characteristics of an Effective Assessment System**

Each assessment should be:

1. linked to standards. Assessments must be tied directly to targeted standards.

2. linked to curriculum and instruction and reflective of the most important content taught. The assessment needs to reflect what students are actually taught. This is why it is critical to align local curriculum to standards.
3. cognitively complex, authentic and integrated. Assessments should model real-world demands. We need to ask for more than simple recall of facts, concepts, principles, and procedures. They should require students to actually apply their knowledge and skills in one or more ways.
4. supportive of self-evaluation and independence. Most educators agree that it is important for students to become self-sufficient learners capable of continually monitoring and contributing to their own learning both inside and outside of school. Cumulative assessments (portfolios and projects) are especially supportive of self-reflection and evaluation.
5. meaningful and flexible. Whenever possible, assessments should not only measure student performance, but also help students further explore and refine important knowledge, skills, and thinking as they apply them in context. In other words, they should be meaningful learning experiences in themselves.
6. responsible to differences in culture, language, cognitive/learning styles and preferred modes of expression. While assessments need some element of standardization (i.e., basic requirements for completion) to ensure comparability of results, they should provide enough diversity in task type and requirements, points of entry, time allowed for completion, opportunities for choice and support to accommodate differences in student cultures, languages, cognitive/learning styles, and modes of representation, aptitudes, and interest.
7. legally defensible. The higher the stakes associated with the assessment results (i.e., Golden State, SAT 9) the greater the emphasis must be on legal defensibility. Validity relates to the degree to which an assessment measures

what it is intended to measure. (The scoring guide or rubric is a key piece of evidence to determine this.) Reliability relates to the consistency of test results, or the degree to which students' assessment performances and the scores on those performances are replicable over time across different circumstances (and scorers).

8. efficient and cost effective. The cost associated with an assessment will depend in large part on the efficiency of the assessment itself, which is influenced by the following factors:
  - a. The amount of content (number and range of standards) that can be covered by the assessment.
  - b. The time and effort needed to administer the assessment (including the support and feedback that must be provided to the students during the assessment).
  - c. The ease and speed with which the assessment can be scored.

Because all assessments have some disadvantages, no one assessment alone can provide a comprehensive view of what students know and can do (e.g. breadth and depth, recall/recognition and application). The ACE and C-TAP models do offer a multiple assessment system which can incorporate career, academic content and schoolwide standards to provide a comprehensive view of student achievement. Specific information on the C-TAP model for assessment is provided in another section of this binder.