Assessment Chapter

of the

Mathematics Framework

for California Public Schools: Kindergarten Through Grade Twelve

Adopted by the California State Board of Education, November 2013

Published by the California Department of Education
Sacramento, 2015

Assessment

Assessment plays a crucial role in delivering high-quality instruction and ensuring that all students learn. For assessment to be effective, teachers need to have sound reasons for selecting and using particular assessment tools. That is, assessment must have a clear purpose in instruction: to support and enhance student learning. Assessment activities should be embedded in instruction and provide opportunities for informative feedback to both students and teachers. A variety of assessment strategies need to be employed, as learning is multi-dimensional and cannot be adequately measured by a single instrument (Suurtamm, Koch, and Arden 2010, 400).

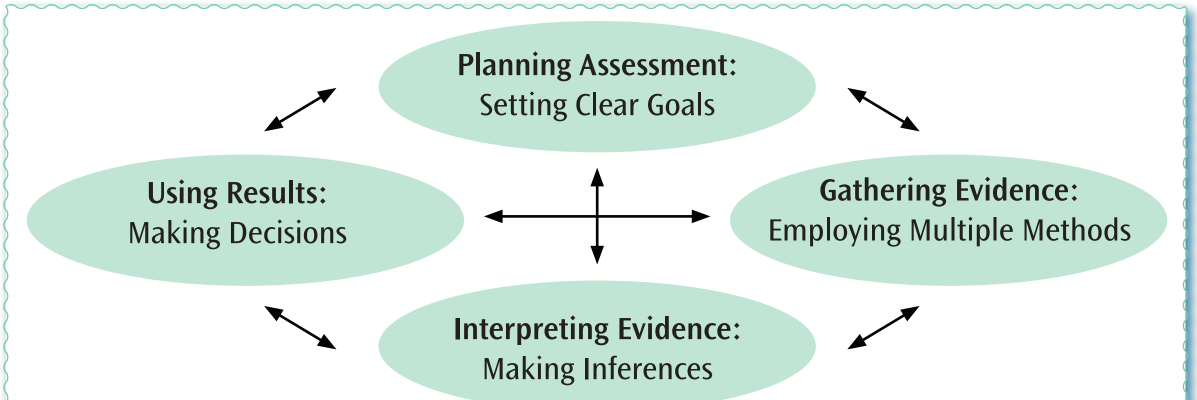
Assessment should be a major component of the learning process. It is essential to instruction because it provides:

- students with frequent and meaningful feedback on their performance;
- teachers with diagnostic tools for gauging students' depth of understanding;
- parents with information about their children's performance in the context of program goals;
- administrators with a means for measuring student achievement.

As students help identify goals for lessons or investigations, they gain greater awareness of what they need to learn and how they will demonstrate that learning. Engaging students in this kind of goal setting can help them reflect on their work, understand the standards to which they are held accountable, and take ownership of their learning.

According to the National Council of Teachers of Mathematics (NCTM), "Assessment is the process of gathering evidence about a student's knowledge of, ability to use, and disposition towards mathematics and of making inferences from the evidence for a variety of purposes" (NCTM 1995, 3). As shown in figure AS-1, the NCTM suggests that there are four interrelated phases of developing assessment and analyzing results.

Figure AS-1. Four Phases of Assessment



Source: NCTM 1995.

The purpose of this chapter is to elucidate some of the key ideas of assessment and provide examples of how to implement them.

Purposes of Assessment

As stated previously, the purpose of assessment should be to support and enhance student learning. A particular assessment may be designed to support the students in an entire school or district, the students in a single classroom, or individual students. Evidence gathered from assessments—regardless of the type of assessment involved—should be used to inform instructional decisions. For example, a teacher may use a mathematics portfolio project to measure students' long-term learning and understanding of the connections among big ideas in a unit, and then use inferences derived from the results to decide how to fill apparent gaps in student understanding before a major summative test. A district or school may use interim assessments (sometimes known as *benchmark assessments*) to track the progress of all grade-five students and then identify schools or classrooms that seem to need the most support in improving student learning. A district may collect statewide testing data and use it to identify populations of students that need support and areas where professional development is needed. If an assessment is implemented without a clearly identified goal for its use or results, then the assessment practice in question should be re-examined; resources may need to be redirected to create an assessment that is more purposeful, or the assessment may need to be eliminated altogether.

At the classroom, department, and possibly school levels, the purpose of assessment of individual students should be more than simply measuring "what students know." Traditional paper-and-pencil and "high-stakes" tests have prompted teachers to emphasize basic, factual information and to provide few opportunities for students to learn how to apply knowledge (Fuchs et al. 1999, 611). Assessment in mathematics must go beyond focusing on how well a student uses a memorized algorithm or procedure and must also elicit, assess, and respond to students' mathematical understandings (NCTM 1995; Suurtamm, Koch, and Arden 2010, 401). This change is essential in light of the Standards for Mathematical Practice, which require students to persevere through solving difficult problems, communicate mathematical thinking, use tools and model with mathematics, use quantities appropriately and attend to precision, and transfer patterns in reasoning and structure to new problems. The focus of assessment must then shift toward assessing content knowledge *and* practices rather than simply assessing content ("what students know how to do"). To help identify learning in both mathematics content and mathematical practice, assessments should ask for variety in what students produce—for example, answers and solutions, arguments and explanations, diagrams and mathematical models.

On a larger scale, assessments may be used to track progress toward long-term learning goals for groups of students or for schools receiving instructional support. Large-scale assessments can help indicate the effectiveness of a professional development program or new instructional materials. Data from statewide summative assessments can be used to identify schools that are performing well in an area or district and those where additional resources can be provided to support improvements in instruction. At the school and district levels, administrators should carefully measure the impact of chosen assessment practices on the classroom; if teachers are constantly under pressure to assess their students, then instruction will often reflect this, and the phenomenon of "teaching to the test" may emerge. Both anecdotal and research evidence show that the undesirable outcome of teaching to the test can and does occur (Fuchs et al. 1999).

Types of Assessment

Current mathematics education literature recognizes two major forms of assessment practices: *formative* and *summative*. The distinction between these types of assessment is based on how they are used, and many forms of assessment can be used both formatively and summatively. Additionally, *diagnostic assessments* are used frequently as tools to place students in courses or identify which students might benefit from an intervention program.

Formative Assessment. Formative assessment is a systematic process to continually gather evidence and provide feedback about learning while instruction is under way. Formative assessment may span a fifteen-minute time period with an individual student, a weeklong unit, or an entire school year. The key feature of formative assessment is that action is taken to close an identified gap in students' learning based on evidence elicited from the assessment practice. As Paul Black and Dylan Wiliam (2001) state in their seminal work on the topic, "assessment becomes 'formative assessment' when the evidence is actually used to adapt the teaching work to meet the needs [of students]" (Black and Wiliam 2001, 2). If an assessment tool is used to gather information and there is no responsive change in instruction to address student misunderstandings, then the tool is not being used formatively. The four phases of assessment come into play with formative assessment, as teachers are often involved in the creation of the assessment tool, the alignment with specific goals, the administration of the tool, and reflection on the results (refer to figure AS-1).

The primary purpose of formative assessment is not merely to audit learning, but to improve it. This is assessment *for* learning rather than assessment *of* learning. Formative assessment is both an instructional tool that teachers and their students "use while learning is occurring" and "an accountability tool to determine if learning has occurred" (National Education Association 2003, 3). In other words, to be *formative*, assessments must inform the decisions that teachers and their students make minute by minute in the classroom. For example, a mid-chapter quiz is usually considered a formative assessment. However, if the result of the quiz is merely recorded in a grade book to serve the purpose of accountability or to certify competence, it cannot be considered a formative assessment. Table AS-1 shows some of the key components of formative assessment in greater detail.

Table AS-1. The Interrelated Dimensions of Formative Assessment

Shared learning targets and criteria for success. A vision of the end point makes the journey possible. Students who have a clear picture of the learning goals and of the criteria for success are likely to have a sense of what they can and should do to make their work measure up to those criteria and goals. They also have some sense of control over their work and are poised to be strategic self-regulators.
Feedback that promotes further learning. The power of formative assessment lies in its double-barreled approach, addressing both cognitive and motivational factors. To be effective, feedback comments should identify what has been done well, point out what still needs improvement, and give guidance on how to make improvements. As part of the overall learning process, teachers should plan for students to have opportunities to respond to comments. Feedback given to any pupil should be about the particular qualities of his or her work and should avoid comparisons with other pupils.
Self-assessment and peer assessment. Many successful innovations have developed self- and peer assessment by pupils as a way of enhancing formative assessment. The main problem for self-assessment is not reliability or trustworthiness; in fact, it is found that pupils are generally honest and reliable in assessing both themselves and one another. They may be too hard on themselves as often as they are too kind. The primary challenge is that pupils can assess themselves only when they have a sufficiently clear picture of the learning targets they are meant to attain. When pupils do acquire such an overview, they become more committed and more effective as learners. Their own assessments become an object of discussion with their teachers and with one another, which promotes learning. Adapted from Black et al. 2004.

Not every formative assessment tool is appropriate for every student, goal, or topic area; therefore, teachers need to differentiate their formative assessment practices based on their experiences of using the tool(s) with their students. Furthermore, formative assessment practices are not necessarily independent of one another; several are often built into the lesson of the day or into the weekly unit.

Summative Assessment. Summative assessment refers to the assessment of learning at a particular point in time; it is meant to summarize a learner's development. Summative assessments frequently take the form of chapter or unit tests, weekly quizzes, or end-of-term tests. In contrast to formative assessment, summative assessment represents the state of a student's skills and knowledge at a particular point in time and is meant to evalute the effectiveness of instruction and a student's learning progress. Such assessments are not necessarily used to inform instruction, but they can be used to measure the effectiveness of an instructional program.

Both summative and formative assessments are essential. However, the crucial distinction between these assessment types is that one aims to determine a student's learning status and the other aims to promote greater learning. Some distinguishing characteristics of formative and summative assessment are provided in table AS-2.

Table AS-2. Characteristics of Formative and Summative Assessment

Formative Assessment (Assessment *for* Learning)	Summative Assessment (Assessment *of* Learning)
Purpose: To improve learning and achievement	**Purpose:** To measure or audit attainment of learning goals
Carried out while learning is in progress—day to day, minute by minute	Carried out from time to time to create snapshots of what has happened
Focused on the learning process and on learning progress	Focused on the products of learning
Viewed as an integral part of the teaching–learning process	Viewed as a separate activity performed after the teaching–learning cycle
Collaborative—Teachers and students know where they are headed, understand the learning needs, and use assessment information as feedback to guide and adapt what they do to meet those needs.	*Teacher directed*—Teachers assign what the students must do and then evaluate how well the students completed the assignment.
Fluid—An ongoing process influenced by student needs and teacher feedback	*Rigid*—An unchanging measure of what the students achieved
Teachers and students adopt the role of intentional learners.	Teachers adopt the role of auditors, and students assume the role of those who are being audited.
Teachers and students use the evidence they gather to make adjustments for continuous improvement.	Teachers use the results to make final "success or failure" decisions about a relatively fixed set of instructional activities.

Adapted from Moss and Brookhart 2009.

Assessment Tools

Many of the assessment tools and strategies listed in this section can be used both formatively and summatively. The list is by no means exhaustive. Furthermore, the tools listed may be administered in a formal way, such as with a checklist of skills for student observation that is filled out for every student throughout a week; or informally, such as with a "ticket-out-the-door" mini-assessment question that is used to gauge student understanding of that day's or week's major concept.

- ***Student observation*** refers to in-classroom observation of students working on mathematics tasks, either independently or in groups. Many teachers already do this by walking around the classroom, actively listening to students, asking questions, directing discourse, and helping students where needed, although they may not see this as a form of assessment. The instantaneous feedback provided to students—suggesting where to go next, recommending questions to ask to gain insight into a problem, correcting computational errors, and so on—results in this practice being a type of formative assessment. Teachers may focus their observations by using checklists that are based on specific skills and concepts.

- ***Graphic organizers*** such as flowcharts and concept maps may be used to assess students' understanding of mathematical concepts and connections between ideas. For instance, a teacher may post several terms in the classroom and ask students to (a) define the terms in their own words; and (b) connect each term to as many other terms as they can, indicating connections with an arrow and providing an explanation of why the terms are related. Teachers can ask students to provide examples of terms or concepts, to explain how and why a certain algorithm or skill works, or describe situations in which a given concept applies.

- ***Student interviews*** can help teachers gain insight into student thinking and guide teachers in providing differentiated instruction. When teachers formally or informally discuss mathematics with students, checking for understanding of concepts or procedures, there is potential to gain a much better understanding of a student's current ability than through the information provided by a paper-and-pencil test. Teachers could use such interviews as a means for assessing student progress on mastering a given standard, and the results of interviews could be factored into grading policies.

- ***Journals and learning logs*** allow students to do mathematical writing that illuminates their current understandings. For example, a teacher may provide each student with a journal—kept in the classroom—that is used for students to solve an "exit problem" of the day. Or students may be asked to explain what they learned that day or what they think the major idea of the lesson was. Such journals have a variety of uses, but teachers should not feel required to grade everything in a math journal; in fact, doing so may diminish its use, as students may feel compelled to write a "correct response." Instead, teachers can periodically read some or all of their students' journals to get feedback on student understanding.

- ***Mathematics portfolios*** are a way to assess students' understanding of important ideas, connections between ideas, procedural knowledge, and the Standards for Mathematical Practice. A project can be explored in groups over several class periods, and a "portfolio" of all the students' relevant work is submitted at the conclusion of the project. Given the nature of the California Common Core State Standards for Mathematics (CA CCSSM) and their emphasis on mathematical *practices*, tools such as portfolios will be necessary to assess students' development as problem solvers and can be used to document students' learning over time.

- ***Self- and peer evaluation*** give students ownership of their learning and provide teachers with insight into students' recognition of their own progress.

- ***Short tests and quizzes*** are used to inform instruction, and small-scale tests and quizzes can be used as formative assessments when integrated as part of a unit or chapter. Such tests and quizzes may involve several different problem types and may or may not contribute to a student's overall course grade. However, if the results of such tests and quizzes are not used to inform future instruction, then these tools are *not* being used formatively.

- ***Performance tasks*** consist of problems or scenarios that require students to think about a problem, encourage students to justify their thinking, and often require students to engage with other students. Administered to individual students or to groups, performance tasks are often complex problem-solving activities that require students to apply prior knowledge in a given situation or to extend current knowledge in new directions. The term *performance task* is broad; it may refer to in-classroom tasks or even to assessment items (see Smarter Balanced Assessment Consortium [Smarter Balanced] 2012c, 31]). Teachers may monitor students' progress on the task and give them immediate feedback as part of a larger formative assessment program.

The CA CCSSM require students to acquire a deep conceptual understanding of mathematics. The introduction of the Standards for Mathematical Practice increases the complexity of gathering evidence to determine student proficiency. Often referred to as *projects*, *oral presentations*, or *written responses to open-ended real-world problems*, performance tasks require a student to demonstrate mathematical learning across several content and practice standards that are considered prerequisite skills for college and career readiness (Smarter Balanced 2012c, 31). Various approaches can be used to determine student proficiency through performance tasks, including rating scales such as rubrics, checklists, and anecdotal records (Burden and Byrd 2010).

On Using Rubrics: A *rubric* is a type of rating scale that allows the teacher to determine mathematical learning along a continuum. By utilizing rubrics, teachers can quantify student learning while focusing upon the predetermined key components of the performance task. Popham (2010) suggests that scoring rubrics have three key features:

1. Evaluative criteria (usually three or four) that indicate the quality of the student's response
2. Descriptions of the qualitative differences in student performance for the evaluative criteria
3. Strategy for scoring, such as whether the performance task will be scored holistically (e.g., a single overall score) or analytically (e.g., points are awarded for each of the performance indicators to provide students with more specific feedback)

Van de Walle (2007) provides an example of a generic four-point rubric (see below) that can be used to sort student responses into two categories before assigning a point designation on a four-point scale. Van de Walle suggests that sharing the rubric with students ahead of time "clearly conveys what is valued" in completing the performance task (Van de Walle 2007, 84).

Scoring with a Four-Point Rubric			
Got It Evidence shows that the student understands the target concept or idea.		**Not Yet** Student shows evidence of major misunderstanding, an incorrect concept or procedure, or failure to engage in the task.	
4 Excellent: Full Accomplishment	**3** Proficient: Substantial Accomplishment	**2** Marginal: Partial Accomplishment	**1** Unsatisfactory: Little Accomplishment
Strategy and execution meet the content, process, and qualitative demands of the task. Communication is judged by effectiveness, not length. May have minor errors.	Could work to full accomplishment with minimal feedback. Errors are minor, so teacher is confident that understanding is adequate to accomplish the objective.	Part of the task is accomplished, but there is a lack of evidence of understanding or evidence of not understanding. Direct input or further teaching is required.	The task is attempted, and some mathematical effort is made. There may be fragments of accomplishment, but there is little or no success.

The Smarter Balanced Assessment Consortium provides examples of rubrics that are based upon the CCSSM, such as the following grade-six problem and scoring rubric that demonstrate student learning for standard **6.EE.5** and mathematical practice standards **MP.1**, **MP.2**, and **MP.4**.

Smarter Balanced Sample Performance Task and Scoring Rubric

Part A

Ana is saving to buy a bicycle that costs $135. She has saved $98 and wants to know how much more money she needs to buy the bicycle.

The equation $135 = x + 98$ models this situation, where x represents the additional amount of money Ana needs to buy the bicycle.

- When substituting for x, which value(s), if any, from the set $\{0, 37, 08, 135, 233\}$ will make the equation true?
- Explain what this means in terms of the amount of money needed and the cost of the bicycle.

Part B

Ana considered buying the $135 bicycle, but then she decided to shop for a different bicycle. She knows the other bicycle she likes will cost more than $150.

This situation can be modeled by the following inequality:

$$x + 98 > 150$$

- Which values, if any, from -250 to 250 will make the inequality true? If more than one value make the inequality true, identify the least and greatest values that make the inequality true.
- Explain what this means in terms of the amount of money needed and the cost of the bicycle.

Sample Top-Score Response:

Part A

The only value in the given set that makes the equation true is 37. This means that Ana will need exactly $37 more to buy the bicycle.

Part B

The values from 53 to 250 will make the inequality true. This means that Ana will need from $53 to $250 to buy the bicycle.

Scoring Rubric: Responses to this item will receive 0–3 points, based on the following descriptions.

3 points: The student shows a thorough understanding of equations and inequalities in a contextual scenario, as well as a thorough understanding of substituting values into equations and inequalities to verify whether they satisfy the equation or inequality. The student offers a correct interpretation of the equality and the inequality in the correct context of the problem. The student correctly states that 37 will satisfy the equation and that the values from 53 to 250 will satisfy the inequality.

2 points: The students shows a thorough understanding of substituting values into equations and inequalities to verify whether they satisfy the equation or inequality, but limited understanding of equations or inequalities in a contextual scenario. The student correctly states that 37 will satisfy the equation and that the values from 53 to 250 will satisfy the inequality, but the student offers an incorrect interpretation of the equality or the inequality in the context of the problem.

1 point: The student shows a limited understanding of substituting values into equations and inequalities to verify whether they satisfy the equation or inequality and demonstrates a limited understanding of equations and inequalities in a contextual scenario. The student correctly states that 37 will satisfy the equation, does not state that the values from 53 to 250 will satisfy the inequality, and offers incorrect interpretations of the equality and the inequality in the context of the problem. **OR** The student correctly states that the values from 53 to 250 will satisfy the inequality, does not state that 37 satisfies the equation, and offers incorrect interpretations of the equality and the inequality in the context of the problem.

0 points: The student shows little or no understanding of equations and inequalities in a contextual scenario and little or no understanding of substituting values into equations and inequalities to verify whether they satisfy the equation or inequality. The student offers incorrect interpretations of the equality and the inequality in the context of the problem, does not state that 37 satisfies the equation, and does not state the values from 53 to 250 will satisfy the equation.

- ***Unit or chapter assessments*** measure student learning of the content and skills in a unit or chapter. Such tests should include items that are linked to specific learning goals, be connected to the CA CCSSM, and pay attention to the Standards for Mathematical Practice. To effectively assess such goals, such tests should include various types of tasks, including multiple choice, selected response (possibly more than one correct response), short answer, and short performance tasks.

- ***Diagnostic assessments*** are often broad in scope, containing a range of topics that are prerequisites for success in a particular unit, class, or grade level. Such assessments may identify specific areas of difficulty that need to be addressed through intervention and can inform the placement of students into intervention programs.

- ***Interim assessments*** can be administered on a relatively frequent basis and are used to measure the incremental learning of students throughout a given period of time. These tests identify specific performance standards students have or have not achieved and often reveal possible reasons why students have not yet progressed in certain areas. Interim assessments are frequently used as formative assessments as well.

- ***State or national assessments*** are large-scale assessments used to gather information about the progress of academic systems and entire bodies of students. (See the section on the Smarter Balanced Assessment Consortium's assessments that appears later in this chapter.)

Thoughts On Grading. Although a classroom grading policy is ultimately a local decision, a message is presented here about the overall purpose and direction of a grading policy. In a chapter titled "The Last Frontier: Tackling the Grading Dilemma" (O'Connor 2007), author Ken O'Connor provides several guidelines for designing grading policies:

- Rather than determining one final grade based only on assessment methods (quizzes, tests, homework, and so forth), teachers should issue grades that are based on and provided for intended learning goals linked to the CA CCSSM.
- Individual achievement should be the primary attribute included in a student's grade. Other aspects such as effort and participation can be graded, but these should not impact measures of achievement.
- Grading should be flexible enough to provide for a sampling of student performance, rather than including every activity and assignment in a grade, and quality assessments with proper recording of student achievement should determine that performance.
- Finally, teachers should discuss assessment with students and involve students in assessment throughout the learning process.

Thoughts on Homework. As with grading policies, whether and how to use homework as an instructional tool and an assessment tool is a local decision. However, if homework is used in a course, it should have clear, standards-based goals that students can achieve on their own. Homework should also promote student ownership of their learning, instill a sense of competence, and be clear and accessible to students. Some reasons for assigning homework include pre-learning of concepts, checking for understanding of classroom work, practice of skills and procedures, and processing of concepts developed in class. Appropriate homework feedback can serve a formative purpose if it provides students and the teacher with direction for learning. As an example, teachers may indicate to students that they should work on problems 1 through 5 first; if these problems are not difficult, then students can move on. However, if a student has difficulty with these first five problems, then that should serve as a warning sign that the student needs to see the teacher for further instruction. Regardless, teachers and administrators should consider a clear purpose for homework as a means for assessment and learning (Van de Walle and Folk 2005).

Smarter Balanced Assessment Consortium, Common Core Assessments

California's participation in the Smarter Balanced Assessment Consortium has resulted in a statewide assessment program designed to measure students' and schools' progress toward meeting the goals of the CA CCSSM for grades three through eight and in grade eleven. Smarter Balanced assessments require students to think critically, solve problems, and show a greater depth of knowledge. They are aligned with the following four claims:

Claim 1	**Concepts and Procedures: Students can explain and apply mathematical concepts and interpret and carry out mathematical procedures with precision and fluency.** This claim addresses procedural skills and the conceptual understanding on which the development of skills depends. It is important to assess students' knowledge of how concepts are linked and why mathematical procedures work the way they do. Central to understanding this claim is making the connection to elements of these mathematical practices as stated in the CA CCSSM: **MP.5**, **MP.6**, **MP.7**, and **MP.8**.
Claim 2	**Problem Solving: Students can solve a range of complex, well-posed problems in pure and applied mathematics, making productive use of knowledge and problem-solving strategies.** Assessment items and tasks focused on Claim 2 include problems in pure mathematics and problems set in context. Problems are presented as items and tasks that are well posed (that is, problem formulation is not necessary) and for which a solution path is not immediately obvious. These problems require students to construct their own solution pathway rather than follow a solution pathway that has been provided for them. Such problems are therefore unstructured, and students will need to select appropriate conceptual and physical tools to solve them.
Claim 3	**Communicating Reasoning: Students can clearly and precisely construct viable arguments to support their own reasoning and to critique the reasoning of others.** Claim 3 refers to a recurring theme in the CA CCSSM content and practice standards: the ability to construct and present a clear, logical, and convincing argument. For older students this may take the form of a rigorous deductive proof based on clearly stated axioms. For younger students this will involve justifications that are less formal. Assessment tasks that address this claim typically present a claim and ask students to provide a justification or counterexample.
Claim 4	**Modeling and Data Analysis: Students can analyze complex, real-world scenarios and can construct and use mathematical models to interpret and solve problems.** Modeling is the bridge between "school math" and "the real world"—a bridge that has been missing from many mathematics curricula and assessments. Modeling is the twin of mathematical literacy, which is the focus of international comparison tests in mathematics given by the Programme for International Student Assessment (PISA). The CA CCSSM feature modeling as both a mathematical practice at all grade levels and a content focus in higher mathematics courses.

Some of the features of the Smarter Balanced assessment program are listed below. Additional information about the assessment program is available at http://www.smarterbalanced.org/ (accessed September 4, 2015).

California now uses a statewide assessment system known as the California Assessment of Student Performance and Progress (CAASPP). Information on the CAASPP system is available at http://www.cde.ca.gov/ta/tg/ca/ (accessed September 4, 2015).

- ***Computer-based testing.*** Schools with the capability to administer tests electronically do so for every student in their purview. Computer-based testing allows for smoother test administration, faster reporting of results, and the utilization of computer-adaptive testing.

- ***Computer-adaptive testing.*** The Smarter Balanced assessments use a system that monitors a student's progress as he or she is taking the assessment and presents the student with harder or easier problems depending on the student's performance on the current item. In this way, the computer system can make adjustments to more accurately assess the student's knowledge and skills.

- ***Varied items.*** The Smarter Balanced tests allow for several types of items that are intended to measure different learning outcomes. For instance, a selected response item may have two correct choices out of four; a student who selects only one of those correct items would indicate a different understanding of a concept than a student who selects both of the correct responses. Constructed-response questions are featured, as well as performance assessment tasks (which include extended-response questions) that measure students' abilities to solve problems and use mathematics in context, thereby measuring students' progress toward employing the mathematical practice standards and demonstrating their knowledge of mathematics content. Finally, the assessments feature technology-enhanced items that aim to provide evidence of mathematical practices.