Chapter 11: Mathematics Assessment in the 21st Century

Table of Contents

Note to reader: The use of the non-binary, singular pronouns *they*, *them*, *their*, *theirs*, *themself*, and *themselves* in this framework is intentional.

Introduction

Assessment is a critical step in the teaching and learning process for students, teachers, administrators, and parents. As a “systematic collection and analysis of information to improve student learning” (Stassen et al., 2001, p. 5), student mathematics assessment is evolving from rote tests of skills to multi-dimensional measures of problem-solving capacity and evidence-based reasoning. This evolution is ongoing in California, as assessments continue to change in order to reflect shifting classroom, school, district, and state priorities. However, as increasingly modern assessments continue to replace traditional tests, all educational assessment should share a common purpose: collecting evidence to enhance student learning, supporting students' development of positive mathematics identities (Aguirre, Mayfield-Ingram, & Martin, 2013).

A comprehensive assessment system consists of summative, interim, and formative assessment. This chapter will start by addressing the need for rethinking the frequency of tests in classrooms that focus only on answer finding. The chapter will discuss the two primary forms of mathematics assessments: **formative assessment** (assessment *for* learning), and **summative assessment** (assessment *of* learning). Each of these will be discussed in detail and describe how they relate to mathematics instruction and learning, with several examples shown.

Broadening Assessment Practices

Important mathematics learning is multidimensional and can be demonstrated through many forms of communication, such as speaking, drawing, writing, and model building. It has long been the practice in mathematics classrooms to assess students' mathematics achievement through narrow tests of procedural knowledge. The knowledge needed for success on such tests is far from the adaptable, critical and analytical thinking needed by students in the modern world. The California Assessment of Student Performance and Progress (CAASPP) has been designed to assess students in responsive and multifaceted ways, capturing their reasoning and problem solving. The irrelevance of previous, more narrow forms of testing resulted in leading employers, such as Google, to declare that they are no longer interested in students' performance on narrow standardized tests that do not reliably predict success in the workplace (Bryant, 2013). Many colleges have eliminated the need for standardized tests for graduate student admissions, and some have eliminated standardized tests for undergraduate student admissions, also citing their lack of usefulness in predicting student success. The sets of skills needed in the modern workforce, even those involving mathematical knowledge, are simply not being accurately assessed by limited tests of mostly procedural or factual knowledge.

Research shows that narrow tests particularly misrepresent the knowledge and understanding of girls and women, leading to inequities in education and employment. In 2012 the team at the Organization for Economic Co-operation and Development (OECD) who conduct the Program for International Student Assessment (PISA) testing conducted a focused analysis on mathematics, with a special report on gender (PISA, 2017a). They found that when taking an individual timed mathematics test, girls achieved at significantly lower levels than boys in 38 countries, despite mathematics achievement being equal in the countries. When the researchers factored in anxiety, the achievement differences disappeared, showing that the under-achievement of girls came from the anxiety provoked by the testing. Further evidence for this was provided by a PISA test of collaborative problem solving (PISA, 2017b). Students were tested individually but they interacted with a computer agent, connecting ideas to help solve complex problems together. In that collaborative assessment, girls out performed boys in all 51 countries. This achievement for girls was matched by another important result. In the collaborative assessment of problem solving there were no differences in the achievement between students from socio-economically advantaged and disadvantaged backgrounds, a result that is very unusual in large scale testing. Considering these two PISA results side by side suggests that girls are disadvantaged in individual tests of mathematics as anxiety reduces their capacity to be successful, but they are enabled in tests that involve collaboration, even with a computer agent. Since the ability to collaborate and to effectively utilize technology are necessary skills in modern workforce environments, modern assessments should, ideally, incorporate these skills.

Narrow tests (tests that focus solely on procedural skills) have also been found to produce racial inequities (https://star.cde.ca.gov/star2012/), and particularly disadvantage language learners (Boaler, 2003). The biased and narrow nature of the tests have been proposed as one of the reasons public perceptions of student ability are often racialized (Martin, 2010). True measurements of learning reflect the need to assess students broadly in order to promote more equitable outcomes as well as more valid assessments of mathematical understanding. Recommendations for equitable teaching and assessing, with clear links between the pursuit of equity and the ways we assess students can be found in Feldman (2019) and DeSilva (2020). A particularly damaging assessment practice to avoid is the use of timed tests to assess speed of mathematical fact retention, as such tests have been found to prompt mathematics anxiety. When anxious, the working memory—the part of the brain needed for reproducing mathematics facts—is compromised. Math anxiety has now been recorded in students as young as five years old (Ramirez, et al, 2013) and timed tests are a major cause of this debilitating, often life-long condition (Boaler, 2014). In recent years, brain researchers have found that the students who are most successful with number problems are those who are using different brain pathways—one that is numerical and symbolic and the other that involves more intuitive and spatial reasoning (Park & Brannon, 2013). Alternative activities can be used that develop mathematics fact fluency through engaging, conceptual, visual activities, instead of anxiety producing, speed tests. Resources for positive and engaging assessment of fact fluency include Boaler (2015), at https://www.youcubed.org/evidence/fluency-without-fear/, (Kling & Bay-Williams, 2013). Inflexible, narrow methods of assessing mathematical competence also disadvantage students with learning differences. The framework of Universal Design for Learning explicitly calls for multi-dimensional assessment practices (Meyer et al., 2014). In mathematics, assessments should be flexible, allowing for multiple means of expression, as well as provide actionable feedback to students (Lambert 2020).

Two Types of Assessment

There are two general types of assessment, formative and summative. **Formative assessment**, commonly referred to as assessment *for* learning, has the goal of providing in-process information to teachers, and students, with regard to learning. Formative assessment is a process teachers and students use during instruction that provides feedback to adjust ongoing teaching moves and learning tactics. It is not a tool or an event, nor a bank of test items or performance tasks. The following definition of formative assessment comes from the *ELA/ELD Framework* (2014):

Figure 8.2. What is Formative Assessment?

What is formative assessment? Formative assessment is a *process* teachers and students use *during* instruction that provides feedback to adjust ongoing teaching moves and learning tactics. It is *not* a tool or an event, nor a bank of test items or performance tasks. Well-supported by research evidence, it improves students' learning in time to achieve intended instructional outcomes. Key features include:

1. ***Clear lesson-learning goals and success criteria***, so students understand what they are aiming for;
2. ***Evidence of learning*** gathered during lessons to determine where students are relative to goals;
3. ***A pedagogical response to evidence, including descriptive feedback***, that supports learning by helping students answer: Where am I going? Where am I now? What are my next steps?
4. ***Peer- and self-assessment*** to strengthen students' learning, efficacy, confidence, and autonomy;
5. ***A collaborative classroom culture*** where students and teachers are partners in learning.

Source

Linquanti, Robert. 2014. *Supporting Formative Assessment for Deeper Learning: A Primer for Policymakers.* Paper prepared for the Formative Assessment for Students and Teachers/State Collaborative on Assessment and Student Standards, 2. Washington, DC: Council of Chief State School Officers.

https://www.cde.ca.gov/ci/rl/cf/documents/elaeldfwchapter8.pdf

Well-supported by ongoing research and evidence, formative assessment improves students' learning in time to achieve intended instructional outcomes (*ELA/ELD Framework*). The CAASPP system encompasses both formative and summative assessment resources, and reflects the work of the Smarter Balanced Assessment Consortium, which further defines formative assessment in the context of the system at https://portal.smarterbalanced.org/library/en/formative-assessment-process.pdf.

Summative assessment, commonly referred to as assessment *of* learning, has the goal of collecting information on a student's achievement *after* learning has occurred. Summative assessment measures include classroom, interim or benchmark assessments, and large-scale summative measures, such as the CAASPP or SAT.

Summative assessments help determine whether students have attained a certain level of competency after a more or less extended period of instruction and learning; such as the end of a unit which may last several weeks, the end of a quarter, or annually (National Research Council [NRC] 2001).

Regardless of the type or purpose of an assessment, teachers should keep in mind that the UDL Principles call for students to be provided "multiple means of action and expression." This can be as simple as allowing students to talk through their solution by pointing and verbalizing, as opposed to strictly writing, or using arrows and circles to highlight particular pieces of evidence in their solution rather than repeating statements in their explanation. Providing a variety of ways for students to showcase what they can do, and what they know, is especially important in mathematics assessments. Aligning assessment with one or more UDL principles can better inform the teacher of what students are learning, and multiple means of representation, whether used to inform formative assessment of daily progress or as a summative display of enduring mathematical understanding, can create a complex and diverse mosaic of student achievement.

An underlying question for teachers as they design, implement, and adapt assessments to be effective for all students is: How can students demonstrate what they know in a variety of ways? Increased use of distance learning has caused a shift in assessment practices which has distinct benefits for students being able to show their understanding in alternative ways. For example, students can video record their thinking related to a task or they can post answers in a live chat or anonymous poll. By considering and planning for the variety of ways in which students can demonstrate their skills and knowledge, they are better able to provide teachers with the information on what they succeed in doing, and where their challenges are.

The main differences between formative and summative assessment are outlined in Table 10.1, which comes from the *ELA/ELD Framework*.
Figure 10.1

Figure 8.3. Key Dimensions of Assessment for Learning and Assessment of Learning

Assessment: A Process of Reasoning from Evidence to Inform Teaching and Learning			
Dimension	**Assessment *for* learning**	**Assessment *of* learning**	
Method	**Formative Assessment Process**	**Classroom Summative/ Interim/Benchmark Assessment***	**Large-Scale Summative Assessment**
Main Purpose	Assist immediate learning (in the moment)	Measure student achievement or progress (may also inform future teaching and learning)	Evaluate educational programs and measure multi-year progress
Focus	Teaching and learning	Measurement	Accountability
Locus	Individual student and classroom learning	Grade level/ department/school	School/district/state
Priority for Instruction	High	Medium	Low
Proximity to Learning	In-the-midst	Middle-distance	Distant
Timing	*During* immediate instruction or sequence of lessons	*After* teaching-learning cycle → *between* units/ periodic	*End* of year/course
Participants	Teacher and Student (T-S/S-S/Self)	Student (may later include T-S in conference)	Student

*Assessment of learning may also be used for formative purposes *if* assessment evidence is used to shape future instruction. Such assessments include weekly quizzes; curriculum embedded within-unit tasks (e.g., oral presentations, writing projects, portfolios) or end-of-unit/culminating tasks; monthly writing samples, reading assessments (e.g., oral reading observation, periodic foundational skills assessments); and student reflections/self-assessments (e.g., rubric self-rating).

Source

Adapted from

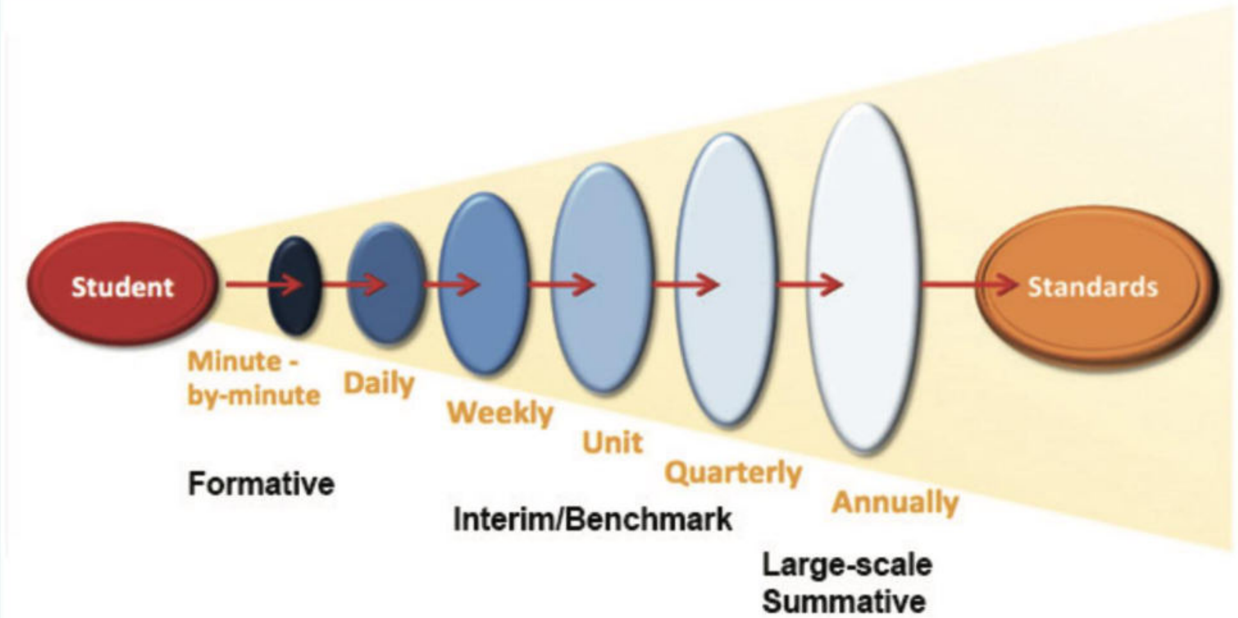
Linquanti, Robert. 2014. *Supporting Formative Assessment for Deeper Learning: A Primer for Policymakers.* Paper prepared for the Formative Assessment for Students and Teachers/State Collaborative on Assessment and Student Standards, 2. Washington, DC: Council of Chief State School Officers.

https://www.cde.ca.gov/ci/rl/cf/documents/elaeldfwchapter8.pdf

The different purposes of assessment cycles are set out in Figure 10.2, from the *ELA/ELD Framework*.

Figure 10.2

Figure 8.4. Assessment Cycles by Purpose



Source

Adapted from

Herman, Joan L., and Margaret Heritage. 2007. *Moving from Piecemeal to Effective Formative Assessment Practice: Moving Pictures on the Road to Student Learning.* Paper presented at the Council of Chief State School Officers Assessment Conference, Nashville, TN.

These are further exemplified in figure 10.3 from the *ELA/ELD Framework*.

Figure 8.5. Types and Uses of Assessments Within Assessment Cycles

Cycle	Methods	Information	Uses/Actions
Short			
Minute-by-Minute	• Observation • Questions (teachers and students) • Instructional tasks • Student discussions • Written work/ representations	• Students' current learning status, relative difficulties and misunderstandings, emerging or partially formed ideas, full understanding	• Keep going, stop and find out more, provide oral feedback to individuals, adjust instructional moves in relation to student learning status (e.g., act on "teachable moments")
Daily Lesson	Planned and placed strategically in the lesson: • Observation • Questions (teachers and students) • Instructional tasks • Student discussions • Written work/ representations • Student self-reflection (e.g., quick write)	• Students' current learning status, relative difficulties and misunderstandings, emerging or partially formed ideas, full understanding	• Continue with planned instruction • Instructional adjustments in this or the next lesson • Find out more • Feedback to class or individual students (oral or written)
Week	• Student discussions and work products • Student self-reflection (e.g., journaling)	• Students' current learning status relative to lesson learning goals (e.g., have students met the goal[s], are they nearly there?)	• Instructional planning for start of new week • Feedback to students (oral or written)

Source: framework.www.cde.ca.gov/ci/rl/cf/documents/elaeldfwchapter8.pdf

Cycle	Methods	Information	Uses/Actions
Medium			
End-of-Unit/Project	• Student work artifacts (e.g., portfolio, writing project, oral presentation) • Use of rubrics • Student self-reflection (e.g., short survey) • Other classroom summative assessments designed by teacher(s)	• Status of student learning relative to unit learning goals	• Grading • Reporting • Teacher reflection on effectiveness of planning and instruction • Teacher grade level/ departmental discussions of student work
Quarterly/Interim/ Benchmark	• Portfolio • Oral reading observation • Test	• Status of achievement of intermediate goals toward meeting standards (results aggregated and disaggregated)	• Making within-year instructional decisions • Monitoring, reporting; grading; same-year adjustments to curriculum programs • Teacher reflection on effectiveness of planning and instruction • Readjusting professional learning priorities and resource decisions
Long			
Annual	• Smarter Balanced Summative Assessment • CELDT • Portfolio • District/school created test	• Status of student achievement with respect to standards (results aggregated and disaggregated)	• Judging students' overall learning • Gauging student, school, district, and state year-to-year progress • Monitoring, reporting and accountability • Classification and placement (e.g., ELs) • Certification • Adjustments to following year's instruction, curriculum, programs • Final grades • Professional learning prioritization and resource decisions • Teacher reflection (individual/grade level/ department) on overall effectiveness of planning and instruction

Formative Assessment

Formative assessment is the collection of evidence to provide day-to-day feedback to students and teachers, so that teachers can adapt their instruction and students become self-aware learners who take responsibility for their learning. Formative assessment is typically classroom-based, and in-sync with instruction, such as analyzing classroom conversations or over-the-shoulder observations of students' diagrams, work, questions, and conversations. There are a number of aspects to effective formative assessment, including embedded formative assessment, rubrics, teacher diagnostic comments, self and peer assessment.

A central goal of formative assessment is encouragement of students to take responsibility for their learning. When teachers communicate to the students where they are now, where they need to be and ways to close the gap between the two places, they provide valuable information to students that enhances their learning. In Black and Wiliam's landmark study (1998 a,b) considering the evidence from hundreds of research studies on assessment, they found that if teachers shifted their practices and used predominantly formative assessment, it would raise the achievement of a country, as measured in international studies, from the middle of the pack to a place in the top five. In addition, Black and Wiliam found that if teachers were to assess students formatively, then the positive impact would outweigh that of other educational initiatives, such as reductions in class size (Black, Harrison, Lee, Marshall, & Wiliam, 2002; Black & Wiliam, 1998a, 1998b). The following table, taken from *Principles to Actions* (NCTM, 2014, p. 56), provides helpful insight into specific teacher and student actions in a formative assessment setting.

Elicit and use evidence of student thinking Teacher and student actions	
What are teachers doing?	What are students doing?
• Identifying what counts as evidence of student progress toward mathematics learning goals. • Eliciting and gathering evidence of student understanding at strategic points during instruction.	• Revealing their mathematical understanding, reasoning, and methods in written work and classroom discourse. • Reflecting on mistakes and misconceptions to improve their mathematical understanding. • Asking questions, responding to, and giving suggestions to support the learning of their classmates.

• Interpreting student thinking to assess mathematical understanding, reasoning, and methods. • Making in-the-moment decisions on how to respond to students with questions and prompts that probe, scaffold, and extend. • Reflecting on evidence of student learning to inform the planning of next instructional steps.	• Assessing and monitoring their own progress toward mathematics learning goals and identifying areas in which they need to improve.

Formative Assessment Lessons

One of the strengths of formative assessment is the flexibility, both in timing and approach, that it affords a classroom teacher. Indeed, one can argue that there are a myriad number of possibilities for teachers to conduct formative assessment throughout a lesson, such as monitoring the types of questions students are asking, the responses students are sharing to questions, and the quality of content in peer conversations. And, though much of this may be unplanned, when formative assessment is intentionally included in a daily lesson plan, the data and analysis are even more effective.

When teachers notice and make sense of student thinking they are given an opportunity to assess formatively (Carpenter et al, 2015; Fernandes, Crespo, & Civil, 2017). The NCTM Principles to Action state that “[e]ffective teaching of mathematics uses evidence of student thinking to assess progress toward mathematical understanding and to adjust instruction continually in ways that support and extend learning” (NCTM, 2014). Complex Instruction is a pedagogical approach, that provides an example of the ways student discussions can give formative information for teachers. Complex Instruction centers upon three principles for creating equity in heterogeneous classrooms through groupwork (Cohen & Lotan, 2014). The first principle involves students developing responsibility for each other, serving as academic and linguistic resources for one another (Cabana, Shreve, & Woodbury, 2014). The second principle involves students working together to complete tasks (Cohen & Lotan, 2014). To realize this principle, teachers must manage equal participation in groups by valuing and highlighting a wide range of abilities and attending to issues of status amongst students (Cohen & Lotan, 2014; Tsu, Lotan & Cossey, 2014). During groupwork, the teacher looks for opportunities to elevate students by highlighting their abilities and contributions to the group, which is referred to as “assigning competence” (Boaler & Staples, 2014). This principle recognizes the fact that group interactions often create status differences between students – and when a teacher perceives that a student has become “low status” in a group, they intervene by publicly praising a mathematical contribution they have made. Underlying these two principles is a third: the implementation of multi-dimensional, group-worthy tasks, which are challenging, open-ended, and require a range of ways of working (Cohen & Lotan, 1997; Banks, 2014). As teachers work to manage heterogeneous groupwork and assign competence they will encounter opportunities to listen to student thinking and to assess formatively. We encourage teachers to plan for student groupings or pairings with language proficiencies in mind. Groupings should be flexible and purposeful and should not be exclusively by proficiency levels, as this can create in-class tracking. English learners (ELs) need opportunities to interact with peers who are native speakers of English, providing ELs with access to language models and authentic opportunities to use their developing language skills.

Snapshot:

A teacher tries a new assessment approach.

Vince is an experienced high school teacher who has been teaching for over 20 years in diverse classrooms, including English learners and learners with special needs. Vince uses a traditional system of testing and grading in his classroom, but recently read about assessment for learning and wondered if the summative assessments he had been using could be used in a formative manner. Instead of giving tests as summative assessments, as he had in previous years, he decided to ask students to answer as many problems as they could.

Before beginning, Vince reviewed the questions as class to be sure everyone understood the directions, or words that may have multiple meanings. This ensured that all students had access to the questions. In consideration of Universal Design for Learning (UDL), he also briefly discussed the multiple modes of expressing their thinking that could be used, including diagrams, words, equations, tables and flowcharts to show steps. When students identified questions which were too difficult, and they could not answer them, he asked them to mark these questions, then use the help of a resource—such as a book or class notes, or translation software—in working out solutions. When students finished the assessment, the work they had done on the marked problems became the work they discussed in class. Vince made sure that as many voices were included in the conversation and that visuals were used. Vince said that the discussions gave him the best information he had ever had on his students' understanding of the mathematics he was teaching.

A rich repository of free lessons supporting teachers in formative assessment are the Classroom Challenges housed at the Mathematics Assessment Resource Service (MARS). Each lesson is structured around an active learning experience for students with a rich task, and teachers are provided with common issues to look for in student responses to questions, as well as samples of, and guidance for, analyzing student work.

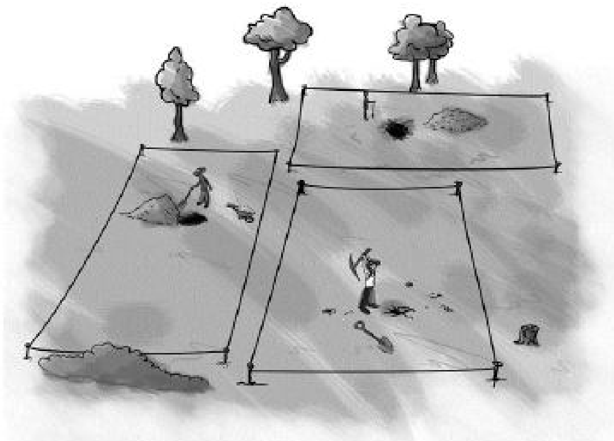
"Maximizing Area: Gold Rush" is a sample grade-seven lesson with following guide to address common student questions. This lesson exemplifies how teachers can adjust their questioning strategies for students based on formative assessment data regarding student misconceptions. (For full lesson: https://www.map.mathshell.org/lessons.php?unit=7300&collection=8)

Gold Rush

In the 19th Century, many prospectors travelled to North America to search for gold.

A man named Dan Jackson owned some land where gold had been found.

Instead of digging for the gold himself, he rented plots of land to the prospectors.



Dan gave each prospector four wooden stakes and a rope measuring exactly 100 meters.

Each prospector had to use the stakes and the rope to mark off a rectangular plot of land.

1. Assuming each prospector would like to have the biggest plot, what should the dimensions of the plot be, once he places his stakes?

   Explain your answer.

2. Read the following statement:

"Join the ropes together!

You can get more land if you work together than if you work separately."

Investigate whether the statement is true for two or more prospectors working together, sharing the plot equally, and still using just four stakes.

Explain your answer.

Common issues:	Suggested questions and prompts:
Does not understand the concept of area and/or perimeter or does not know how to find the area and perimeter of a rectangle	• What does the length of the rope given to a prospector measure? • How could you measure the amount of land enclosed by the rope? • How do you find the area of a rectangle? • How do you find the perimeter of a rectangle?
Calculates the total amount of land, but not the amount of land for each prospector (Q2)	• You've worked out the total area of land for both/all the prospectors; how much land will each prospector get?
Emphasizes only the human impact of sharing the land (Q2) For example: The student states that when two people share they can help each other out. Or: The student states that when sharing the land people are more likely to steal from each other.	• Now investigate if combining ropes affects how much land each prospector gets.
Does not investigate any or very few rectangles For example: The student draws just one rectangle and calculates its area (Q1).	• Now investigate the area of several different rectangles with the same perimeter, but different dimensions.
Works unsystematically	• How can you now organize your work? • How do you know for sure your answer is the best option?
Presents work poorly For example: The student presents the work as a series of unexplained numbers and/or calculations.	• Would someone unfamiliar with this work understand your method?
Only investigates two prospectors sharing land	• Suppose 3/4/5 prospectors share land. What area of land would each prospector get?

Rubrics

Although rubrics are often used by teachers as a tool to evaluate summative work and identify more reliable scores when grading student work, rubrics lend themselves to the formative assessment process because they can provide students with a clear set of expectations to achieve as they learn, and ultimately sure of a success criteria for summative assessment. A rubric can set out for students the mathematics they are learning, enabling students to develop self-awareness and to reflect on their own progress. Oftentimes students can carefully answer questions in lessons, but have difficulty in connecting their learning to the broader mathematical landscape. In association with a rubric, teachers can enact important methods such as self and peer assessment and the provision of comments guiding students to making important connections to other areas of their mathematical knowledge. In creating rubrics, teachers should be mindful of the variety of ways in which students can demonstrate their knowledge. Similar to providing multiple modes of engagement for students during instruction, teachers can provide multiple options for students to showcase their skills and knowledge. For example, teachers can provide colored tape so students can make tape diagrams rather than drawing each and shading. Or teachers can use a camera to take a sequence of images to document students' work while using manipulatives, such as integer chips, to solve a problem, thus saving much copying and drawing.

As seen in the rubric examples provided below, the criteria can focus on the mathematical practices, mathematical content, or both. The following two rubrics, created at the Stanford Center for Assessment Learning and Equity (SCALE), communicate the mathematical practices in a form that students can use to monitor their own progress and learning: http://performanceassessmentresourcebank.org/system/files/PARB%20CC%20BY%204.0%20SCALE%20Math%20PA%20Rubric%20Gr3-12%202016.pdf

SCALE MATH PERFORMANCE ASSESSMENT RUBRIC, GRADES 3-12

Practice	Not Yet	Approaches	Achieves	Masters
Make sense of problems and persevere in solving them	I need assistance from my teacher to understand what the problem or question asks me to do. I am unsure how to connect this problem or question to what I already know. I am still working to organize the information in this problem or question.	I have a partial understanding of what a problem or question asks me to do. I am working on this to make the connection stronger. I show a partial connection between this question and what I already know. I am working on this to make the connection stronger. I organized some of the information in this question or problem but missed some important information.	I explain questions and problems in my own words. I relate questions and problems to similar things I have seen before. I organize given information before attempting to solve. I check to make sure that my final solution makes sense and is reasonable.	Achieves, and also: My work includes a reflection of how I monitored myself while I was working and adjusted my plan when necessary.
Reason abstractly and quantitatively	I am still working to translate between my math work (symbols, calculations) and real world situations. I currently do this with the assistance of my teacher.	I show and explain what some of my math work (symbols, calculations) means in real life contexts.	I show and explain what all or most of my math work (symbols, calculations) means in real life contexts. I pay attention to the meaning of quantities, not just how to compute them.	Achieves, and also: I describe my solution and any limitations in terms of the real world context described within the problem.

SCALE
Stanford Center for Assessment, Learning, & Equity

PERFORMANCE ASSESSMENT RESOURCE BANK

Math Performance Assessment Rubric (Grades 9-12)

The ability to reason, problem-solve, develop sound arguments or decisions, and create new ideas by using appropriate sources and applying the knowledge and skills of a discipline.

	EMERGING	E/D	DEVELOPING	D/P	PROFICIENT	P/A	ADVANCED
PROBLEM SOLVING *What is the evidence that the student understands the problem and the mathematical strategies that can be used to arrive at a solution?*	• Does not provide a model • Ignores given constraints • Uses few, if any, problem solving strategies		• Creates a limited model to simplify a complicated situation • Attends to some of the given constraints • Uses inappropriate or inefficient problem solving strategies		• Creates a model to simplify a complicated situation • Analyzes all given constraints, goals and definitions • Uses appropriate problem solving strategies		• Creates a model to simplify a complicated situation and identifies limitations of model • Analyzes all given constraints, goals and definitions and implied assumptions • Uses novel problem solving strategies and/or strategic use of tools
REASONING AND PROOF *What is the evidence that the student can apply mathematical reasoning/procedures in an accurate and complete manner?*	• Provides incorrect solutions without justifications • Results are not interpreted in terms of context		• Provides partially correct solutions or correct solution without logic or justification • Results are interpreted partially or incorrectly in terms of context		• Constructs logical, correct, complete solution • Results are interpreted correctly in terms of context		• Constructs logical, correct, complete solution with justifications • Interprets results correctly in terms of context, indicating the domain to which a solution applies • (Monitors for reasonableness, identifies sources of error, and adapts appropriately)
CONNECTIONS *What is the evidence that the student understands the relationships between the concepts, procedures, and/or real-world applications inherent in the problem?*	• Little or no evidence of applying previous math knowledge to given problem		• Applies previous math knowledge to given problem but may include reasoning or procedural errors		• Applies and extends math previous knowledge correctly to given problem		• Applies and extends previous knowledge correctly to given problem; makes appropriate use of derived results • (Identifies and generalizes the underlying mathematical structures of the given problem to other seemingly unrelated problems or applications)
COMMUNICATION AND REPRESENTATION *What is the evidence that the student can communicate mathematical ideas to others?*	• Uses representations (diagrams, tables, graphs, formulas) in ways that confuse the audience • Uses incorrect definitions or inaccurate representations		• Uses representations (diagrams, tables, graphs, formulas), though correct, do not help the audience follow the chain of reasoning; extraneous representations may be included • Uses imprecise definitions or incomplete representations with missing units of measure or labeled axes		• Uses multiple representations (diagrams, tables, graphs, formulas) to help the audience follow the chain of reasoning • With few exceptions, uses precise definitions and accurate representations including units of measure and labeled axes		• Uses multiple representations (diagrams, tables, graphs, formula) and key explanations to enhance the audience's understanding of the solution; only relevant representations are included • Uses precise definitions and accurate representations including units of measure and labeled axes; uses formal notation

http://performanceassessmentresourcebank.org/system/files/PARB%20CC%20BY%204.0%20SCALE%20Math%20Performance%20Assessment%20Rubric%20Gr%209-12%202013.pdf

The following rubric from the 2013 *Mathematics Framework* provides criteria for a performance task:

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

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.	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.

An engaging mathematics task that draws from mathematical and scientific understanding, is provided in the *Science Framework* for California. The task is accompanied by a rubric, that the teacher, Mr. A, used to assess the students' work:

Assessment Snapshot 9.6: Mathematical Thinking for Early Elementary

Mr. A's kindergarten class was conducting an investigation when they realized that they needed to use **mathematical thinking [SEP-5]**. Mr. A's class received a package of silkworm eggs and was amazed how they all hatched on almost the same day. One student asked how quickly they would grow and another wondered how big they would get. The students decided that they would like to track the **growth [CCC 7]** of their silkworms and measure them daily. Mr. A wanted the students to come up with a way to answer the question "How **big [CCC-3]** are they today?" through a visual display of their measurement data. The students needed to find a way to summarize all their measurements using a graphical display. Mr. A was guided by research about the different developmental levels in understanding how to display data (table 9.4).

Table 9.4. Developmental Levels of the Ability to Display Data

Level	Descriptor
6	Create and use data representations to notice trends, patterns, and be able to recognize outliers.
5	Create and use data representations that recognize scale as well as trends or patterns in data.
4	Represent data using groups of similar values and apply consistent scale to the groups.
3	Represent data using groups of similar values (though groups are inconsistent).
2	Identify the quantity of interest, but only consider each case as an individual without grouping data together.
1	Group data in ways that don't relate to the problem of interest.

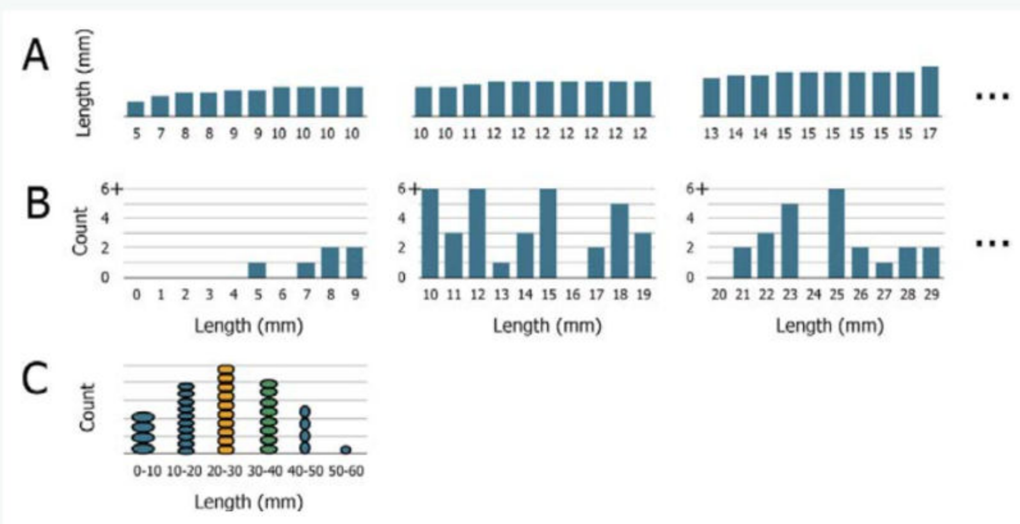
Source: Adapted from Wilson, Ayers, and Schwartz 2013.

Assessment Snapshot 9.6: Mathematical Thinking for Early Elementary

One group ordered each of the 261 measurements by magnitude, making a bar for each worm. The display used a full 5 feet of wall space! (figure 9.A; level 2 on table 9.4). Another group made a bar graph with a bin size of just 1 mm per bin, which led to 50 different bars (figure 9.13B; level 4 on table 9.4). Also, this group's vertical axis only extended to six worms at the top of the paper, so bars with more than six worms were cut off. A third group created a more traditional bar graph with measurements placed into bins. Rather than using bars, the group used circles stacked one on top of the other. Unfortunately, different students drew the circles for each bin, and they were not the same size and therefore not comparable (figure 9.13C; level 3 on table 9.4).

Mr. A led a discussion about which representations were most useful for understanding silkworm growth. Mr. A recognized that each representation was at a different developmental level and used that understanding to highlight different concepts with different students (grouping versus consistent grouping, for example). As students **examined the graphs [SEP-4]** with better understanding of what they represented, they noticed a **pattern [CCC-1]** that there were more "medium-sized" silkworms and fewer short or long ones (level 5 on table 9.4), which allowed Mr. A to introduce the concept of variability. Students began to ask questions about why some silkworms were growing so much faster than others. Mr. A's targeted guidance about how to represent data helped elevate the scientific discussion.

Figure 9.13. Facsimiles of Student-Created Representations of Silkworm Length Data



Groups A and B continue off to the right with additional pages.

Source: Adapted from Lehrer 2011.

Long description of Figure 9.13.

Assessment Snapshot 9.6: Mathematical Thinking for Early Elementary

Commentary:

SEPs. The emphasis of the rubric is on the ability to count and recognize similar values, examples of using **mathematical thinking [SEP-5]** at the primary level.

DCIs. While the activity supports the DCIs that plants and animals have unique and diverse life cycles (LS1.B) and that individuals can vary in traits (LS3.B), the task does not assess student understanding of these DCIs.

CCCs. Students could not complete this task without attention to **scale and quantity [CCC-3]**, including the use of standard units to measure length. The rubric in table 9.4 emphasizes student ability to recognize **patterns [CCC-1]** as they create their data representations.

Resources:

Based on NRC 2014.

Source: www.cde.ca.gov/ci/sc/cf/documents/scifwchapter9.pdf

Some teachers choose to give rubrics to students based around one mathematical area or standard, these are sometimes referred to as "single point rubrics":

Ways I could improve	Criteria	I have shown this in:
	I approach problems in different ways - using drawings, words, and color coding to connect ideas.	

Source: https://www.cultofpedagogy.com/single-point-rubric/

Single point rubrics provide a way for teachers to focus on something important and to give diagnostic comments and diagnostic teacher feedback (see next section) on a particularly important area of work.

Teacher Diagnostic Comments

Assessment for learning communicates to students where they are in their mathematical pathway and, often, how they may move forward. One way to communicate feedback is by sharing grades students have earned, but grades do not give feedback to students about ways to improve. Teacher diagnostic comments are an important part of this communication and allow teachers to share with students their knowledge of ways to improve or build upon their thinking. Different researchers have compared the impact of grades with diagnostic feedback.

Elawar and Corno, for example, contrasted the ways students responded to mathematics homework in sixth grade, with half of the students receiving grades and the other half receiving diagnostic comments without a grade (Elawar & Corno, 1985). The students receiving comments learned twice as fast as the control group, the achievement gap between male and female students disappeared, and student attitudes improved.

Ruth Butler also contrasted students who were given grades for classwork with those who were given diagnostic feedback and no grades (Butler, 1987, 1988). Similar to Corno and Elawar, the students who received diagnostic comments achieved at significantly higher levels. In Butler's study a third condition was added, when students received grades *and* comments—combining both forms of feedback. However, this showed that the students who received grades only and those who received grades and comments scored at similar levels, and the group that achieved at significantly higher levels was the comment-only group. When students received a grade and a comment, they appeared only to focus on the grade. Butler found that both high-achieving (the top 25-percent grade point average) and low-achieving (the bottom 25-percent grade point average) fifth and sixth graders suffered deficits in performance and motivation in both graded conditions, compared with the students who received only diagnostic comments.

Pulfrey, Buchs, and Butera (2011) followed up on Butler's study, replicating her finding—showing again that students who received grades as well as students who received grades and comments both underperformed and developed less motivation than students who received only comments. They also found that students needed only to *think* they were working for a grade to lose motivation, resulting in lower levels of achievement.

Teachers, quite rightly, worry about the extra time that diagnostic feedback can take but diagnostic comments can be given occasionally, instead of frequent grading of class or homework, providing students with insights that can propel them onto paths of higher achievement. A teacher giving comments to students once a week is more useful than frequent grades and test scores. Many learning management systems (LMS) allow teachers to give students verbal feedback on their work. The following example of student work comes from the Interactive Mathematics Program (IMP): The High Dive Problem (https://stephanheuer.wordpress.com/2008/05/09/math-high-dive-unit-problem/).

The teacher comments, in green, are an example of teacher diagnostic comments—some of which are encouraging, some questioning, and some guiding (Boaler, Dance, Woodbury, 2018).

High Dive – Checkpoint 1 **Name**

While on a road trip with your family, you stop for lunch in a small town that has a Ferris wheel. This Ferris wheel has a radius of 30 feet, the center of the wheel is 35 feet above the ground, and the wheel completes one full rotation in 90 seconds. (The Ferris wheel still rotates counter clockwise.)

You want to impress your family by telling them how high off the ground you are at certain times. To convince your family of your expertise you justify your solutions by including labeled diagrams and organized work.

1. What is your height off the ground 18 seconds after you pass the 3:00 position.

$360°/90_{sec} = 4°/sec$ X = Opposite } Good strategy for starting the problem..

$4 * 18 = 72°$ angle

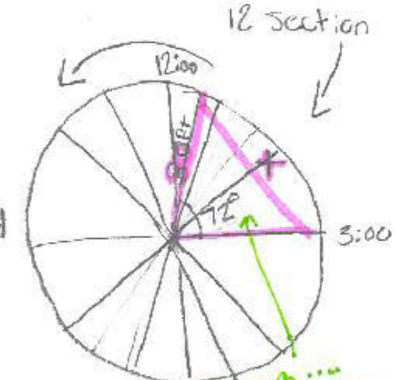
$28.53 + 35 = 63.53$ ft off the ground

$*30$ $\sin(72) = \frac{X}{30}$ $*30$

$30 * \sin(72) = X$

$28.53 = X$

→ what does this number represent?



hmmm... doesn't look like a right triangle

2. What is your height off the ground 35 seconds after you pass the 3:00 position.

$360°/90_{sec} = 4°/sec$

* Trig works with angles bigger than 90° because of inversion *

??? what does this mean

$4° * 35_{sec} = 140°$

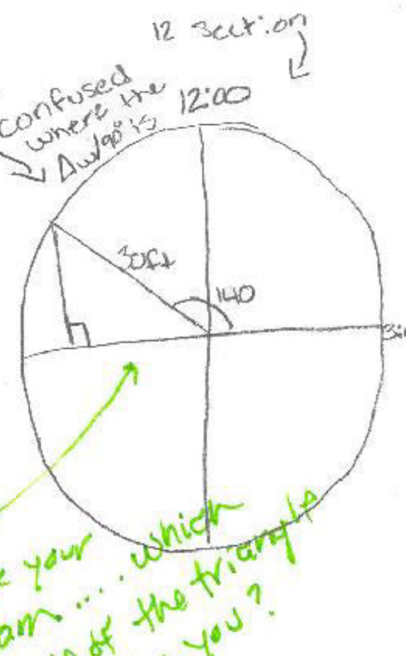
$*30$ $\sin(140) = \frac{X}{30}$ $*30$

$19.28 + 35 = 54.28$ ft. Off the ground

$30 * \sin(140) = X$

$19.28 = X$

Thank you for justifying your work!!



confused where the 90° is

I like your diagram.... which side of the triangle helps you?

Self- and Peer Assessment

The two main strategies for helping students become aware of the mathematics they are learning and their broader learning pathways are self- and peer assessment. In self-assessment, students are given clear statements of the mathematical content and practices they are learning, which they use to think about what they have learned and what they still need to work on. The statements could communicate mathematics content such as, "I understand the difference between mean and median and when each should be used," as well as mathematical practices such as, "I have learned to persist with problems and keep going even when they are difficult." If students start each unit of work with clear statements about the mathematics they are going to learn, they start to focus on the bigger landscape of their learning journeys—they learn what is important, as well as what they need to work on to improve. Studies have found that when students are asked to rate their understanding of their work through self-assessment, they are incredibly accurate at assessing their own understanding, and they do not over- or underestimate it (Black et al., 2002).

Self-assessment can be developed at different degrees of granularity. Teachers could give students the mathematics in a lesson or show them the mathematics from a longer period of time, such as a unit or even a whole term or semester. In addition to receiving the criteria, students need to be given time to reflect upon their learning, which they can do during a lesson, at the end of a lesson, or even at home. The following example is an algebra content self-assessment (Boaler, 2016):

Algebra 1 Self-Assessment

Unit 1 – Linear Equations and Inequalities

- I can solve a linear equation in one variable.
- I can solve a linear inequality in one variable.
- I can solve formulas for a specified variable.
- I can solve an absolute value equation in one variable.
- I can solve and graph a compound inequality in one variable.
- I can solve an absolute value inequality in one variable.

Unit 2 – Representing Relationships Mathematically

- I can use and interpret units when solving formulas.
- I can perform unit conversions.
- I can identify parts of an expression.
- I can write the equation or inequality in one variable that best models the problem.
- I can write the equation in two variables that best model the problem.
- I can state the appropriate values that could be substituted into an equation and defend my choice.
- I can interpret solutions in the context of the situation modeled and decide if they are reasonable.
- I can graph equations on coordinate axes with appropriate labels and scales.
- I can verify that any point on a graph will result in a true equation when their coordinates are substituted into the equation.
- I can compare properties of two functions graphically, in table form, and algebraically.

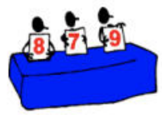
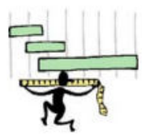
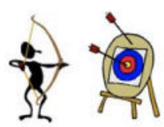
Unit 3 – Understanding Functions

- I can determine if a graph, table, or set of ordered pairs represents a function.
- I can decode function notation and explain how the output of a function is matched to its input.
- I can convert a list of numbers (a sequence) into a function by making the whole numbers the inputs and the elements of the sequence the outputs.

Peer assessment is a similar strategy to self-assessment, as it also involves giving students clear criteria for assessment, but they use it to assess each other's work rather than their own. When students assess each other's work, they gain additional opportunities to become aware of the mathematics they are learning and need to learn. Peer assessment has been shown to be highly effective, in part because students are often much more open to hearing criticism or a suggestion for change from another student, and peers usually communicate in ways that are easily understood by each other (Black et al., 2002).

One method of peer assessment is the identification of "two stars and a wish." Students are asked to look at their peers' work and, with or without criteria, to select two things done well and one area to improve on. When students are given information that communicates clearly what they are learning, and they are asked, at frequent intervals, to reflect on their learning, they develop responsibility for their own learning. Some people refer to this as inviting students into the guild—giving students the powerful knowledge—knowledge that usually only teachers hold—which allows them to take charge of their learning.

Note a self-assessment example that focuses on mathematical practices:

Standard for Mathematical Practice	Student Friendly Language
1. Make sense of problems and persevere in solving them. 	• I can try many times to understand and solve a math problem.
2. Reason abstractly and quantitatively. 	• I can think about the math problem in my head, first.
3. Construct viable arguments and critique the reasoning of others. 	• I can make a plan, called a strategy, to solve the problem and discuss other students' strategies too.
4. Model with mathematics. 	• I can use math symbols and numbers to solve the problem.
5. Use appropriate tools strategically. 	• I can use math tools, pictures, drawings, and objects to solve the problem.
6. Attend to precision. 	• I can check to see if my strategy and calculations are correct.
7. Look for and make use of structure 	• I can use what I already know about math to solve the problem.
8. Look for and express regularity in repeated reasoning. 	• I can use a strategy that I used to solve another math problem.

Source: Rhode Island Department of Education: https://www.ride.ri.gov/Portals/0/Uploads/Documents/Instruction-and-Assessment-World-Class-Standards/Transition/EIA-CCSS/ScarpelliD-MP_ICanStatements.pdf

Standards-Based Grading

Standards-based grading is a way to bring some of the very valuable aspects of formative assessment into summative assessments. This method of assessment shifts the focus from a fixed measure based on a score or a test result to a reflection of the standards students are working towards. Standards-based grading breaks content down into learning targets, each of which is a teachable concept for which students may demonstrate proficiency. Instead of offering partial credit for incorrect responses, students are provided feedback and opportunity to re-assess standards they do not meet in their first attempt. Teachers can then track and provide feedback based on students' work in relation to each learning target. The first table below lists an example of learning targets on a standards-based report card. The second table demonstrates the difference in grading scales. The CAASPP uses a scale which is based on proficiency.

[Note: The following examples will be adapted rather than taken verbatim from websites.]

Traditional Report Card	
Class	**Q1**
Mathematics	95% = A

Standards-based Report Card	
Class/Standards	**Q1**
Mathematics	3
I can define a number sentence	2
I can solve number sentences that have brackets	2
I can solve number sentences that have braces	3
I can create number patterns using two rules	3
I can estimate the answers of number sentences	2
I can find the sum of two 2-digit numbers	3
I can find the difference of two 2-digit numbers	2
I can find the product of two 2-digit numbers	2
I can find the quotient of two 2-digit numbers	3

https://www.teacherease.com/standards-based-grading.aspx

A different example from a kindergarten teacher is shown here:

Kindergarten Mathematics (continued)	**Q1**	**Q2**	**Q3**	**Q4**
Number and Operations in Base Ten				
I work with numbers 11–19 to show ten ones and some further ones.				
Measurement and Data				
I describe, compare, and classify objects and count the number in each category.				
Geometry				
I identify and describe flat and 3D shapes.				
I compare, create, and compose shapes.				
Kindergarten Music	**Q1**	**Q2**	**Q3**	**Q4**
I understand musical concepts.				
I demonstrate knowledge of musical skills.				
I participate appropriately.				
Kindergarten Physical Education	**Q1**	**Q2**	**Q3**	**Q4**
I demonstrate knowledge of P.E. skills				
I demonstrate sportsmanship and cooperate with classmates.				
I actively participate during physical activities.				
Kindergarten Science	**Q1**	**Q2**	**Q3**	**Q4**
I demonstrate an understanding of scientific content and concepts.				
Kindergarten Social Studies	**Q1**	**Q2**	**Q3**	**Q4**
I demonstrate an understanding of social studies content and concepts.				

http://www.isbestandardsbasedreporting.com/report-card-examples.html

Below, note further examples from Saddleback Valley Unified School District:

MATHEMATICS **Effort**			
Ratios and Proportional Relationships			
Understands ratio concepts and uses ratio reasoning to solve problems			

The Number System			
Applies and extends previous understandings of multiplication and division to divide fractions by fractions			
Applies and extends previous understandings of numbers to the system of rational numbers			
Expressions and Equations			
Applies and extends previous understandings of arithmetic to algebraic expressions			
Solves one-variable equations and inequalities			
Represents and analyzes quantitative relationships between dependent and independent variables			
Geometry			
Solves real-world and mathematical problems involving area, surface area, and volume			
Statistics and Probability			
Develops understanding of statistical variability			
Summarizes and describes distributions			

MATHEMATICS Effort			
Operations and Algebraic Thinking			
Represents and solves problems involving addition and subtraction			
Adds and subtracts fluently within 20			
Works with equal groups of objects to gain foundations for multiplication			
Numbers and Operations in Base Ten			
Understands and applies place value concepts			
Uses place value understanding and properties of operations to add and subtract			
Measurement and Data			
Measures and estimates lengths in standard units			
Relates addition and subtraction to length			
Works with time and money			
Represents and interprets data			
Geometry			
Reasons with shapes and their attributes			

David Douglas School:

MATHEMATICS	S1	S2
Read, write, compare, and round decimals to thousandths. Convert metric measurements. **NBT.3**, NBT.1-4, MD.1		
Fluently multiply multi-digit whole numbers using the standard algorithm. Convert customary measurements. **NBT.5, MD.1**		
Solve multi-digit (up to 4 digit by 2 digit) whole number division problems using various strategies. **NBT.6**		
Add, subtract, multiply, and divide decimals to the hundredths place using various strategies. **NBT.7**		
Solve real-world and mathematical problems involving addition and subtraction of fractions including unlike denominators. Make line plots with fractional units. **NF.2**, NF.1, MD.2		
Solve real-world and mathematical problems involving multiplication of fractions and mixed numbers, including area of rectangles. **NF.6**, NF.4, NF.5		
Solve real-world and mathematical problems involving division of fractions by whole numbers (1/4 ÷ 7) and division of whole numbers by fractions (3 ÷ 1/2). Interpret a fraction as division. **NF.7**, NF.3		
Solve real-world and mathematical problems involving volume by using addition and multiplication strategies and applying the formulas. **MD.5**, MD.3-5		
Solve real-world and mathematical problems by graphing points, including numeral patterns, on the coordinate plane. **G.2**, G.1, OA.3		

Loma Prieta Joint Union School District:

MATHEMATICS	T1	T2	T3
Operations and Algebraic Thinking			
Use Operations with Whole Numbers to Solve Problems			
Gain Familiarity with Factors and Multiples			
Generalize and Analyze Problems			
Number and Operation Base 10			
Understand Place Value for Multi-Digit Whole Numbers			
Use Place Value Understanding and Properties of Operations to Perform Multi-Digit Arithmetic			
Number Operations and Fractions			
Understanding of Fraction Equivalence and Ordering			
Build Fractions from Unit Fractions			
Understand Decimal Notation for Fractions			
Measurement Data			
Solve Problems Involving Measurement and Conversion			
Represent and Interpret Data			
Geometry			

Draw, Identify, and Utilize Lines and Angles			

Standards-based grading can be reported to districts, parents and others in the form of standards achieved and not associated with letter grades. Alternatively, teachers can develop structures and methods that turn standards-based-grading results into letter grades if required. These systems could be tied to the percentage of standards mastered, the number of standards at different levels, or tied to mastery of key learning outcomes and some amounts of additional material. One key benefit of using standards-based grading is that it includes a lot more information on what students actually know. When it includes opportunities for reassessment, and students working with feedback to improve their results, it also encourages important growth mindset messages. Researchers have considered parents' responses to a shift to standards-based grading, finding that parents are supportive of standards-based grading, as an alternative to traditional grading (Brookhart et al, 2016). Other research studies have shown that standards-based grading improves student engagement and achievement (Iamarino, 2014; Townsley et al, 2016; Selbach-Allen et al, 2020).

On a final note, since standards-based grading is based on students' meeting of learning targets, grade reports function differently. Test and quiz scores, for example, as percentages are often averaged and translated to letter grades in a traditional system whereas, in a standards-based grades system, mastery of topics is evidenced and communicated over time and in multiple ways. At early points in the year, it should not be expected that students would have mastered all, or even a significant number, of learning targets and grade reports would reflect this progression. Schools should provide clear and consistent messaging regarding standards-based grading systems to help parents, and students, understand report cards.

In traditional grading systems points are often offered for participation, attendance, behavior and homework completion. These measures often bring inequity into the grading system as students outside circumstances will impact these aspects of their grade. The final grade becomes more about behaviors than learning. While standards-based grading is not a panacea to fix inequities in assessments it ties grades and assessment directly back to demonstrated knowledge rather than behaviors that may not reflect student's actual learning.

Effective Assessment Strategies for English Learners

Recognizing the interdependency of disciplinary language and content, we recommend that teachers formatively assess students' use of language in the context of mathematical reasoning over time. At the outset of a unit, students would likely use more exploratory language including everyday language, and over the course of the units, students would add to their repertoire the more standard, less ambiguous, form of mathematical conventions and agreements. One of several Mathematical Language Routines (https://ell.stanford.edu/sites/default/files/u6232/ULSCALE_ToA_Principles_MLRs__Final_v2.0_030217.pdf) that has been developed is called "Collect and Display" (Zwiers and colleagues, 2017, pg. 11) where the teachers listen to students' use of language, then they display the collection of terms they heard and this becomes a language resource for the class. Such a record is useful as it shows the development of language over the time.

Teachers should also provide rubrics, including a discussion of key academic vocabulary, so that the criteria for success is clear to students. Because rubrics can be used to conduct self- and peer assessments (in addition to assessment by the teacher), it can be useful for teachers to provide language instruction, including frames for collaborative criteria chats, if key terms are expected in students' explanations.

For culminating assessments, teachers should do an analysis of the language demands prior to administering the assessments, and backwards planning, guided by the following questions: "What opportunities are provided for students to explain and elaborate their reasoning? Prior to the assessment, have students have had sufficient opportunities to practice using the kind of language that is expected to demonstrate their mathematical reasoning? Have students received feedback and a chance to apply that feedback to their work?

EXAMPLE: In a unit test, suppose students are asked to explain how they know that a linear systems of equations has no solutions. Throughout the instructional unit, students should have opportunities to generate and refine such explanations, working on specific cases but also building up to the language of generalization over time. Students should examine examples of explanations that include visuals of parallel lines, along with a focus on the slopes of the given lines in this case. Using language for complex ideas is an attainable goal for English learners, but only if there is a thoughtful ramp up and support throughout the instruction.

Feedback on student explanations on assessments should follow the same principles of high-quality feedback for English learner students: feedback should acknowledge what was done correctly, ask clarifying questions, and give students an opportunity to revise their work.

As teachers continue to collect formative data about students language, they can act on that data by assessing growth over time, adjust instruction, and consider possible flexible groupings to provide more targeted support.

Teachers may consider the following assessment modifications appropriate for newcomers and ELs who are in the process of acquiring English:

- Allow students to answer verbally rather than in writing, or some combination.
- Consider chunking longer assessments into smaller parts
- Enlist a qualified bilingual professional to help with assessments.
- Consider offering ELs to demonstrate progress in group assessments.
- Allow students to give responses in multiple formats and with the support of manipulatives.
- Accept responses in the students’ native language if translation support systems exist in the school

- Allow ELs to use a bilingual dictionary or translation software to support their language learning.

Summative Assessment

Summative assessment is assessment of learning. Summative assessments typically are given at the end of a learning cycle in order to ascertain students' acquisition of knowledge and skills in the subject. On a classroom level, exams, quizzes, worksheets, and homework have traditionally been used as summative measures of learning for particular units or chapters. Summative assessments have the potential to be anxiety-inducing for students, so some best practices should be implemented to minimize damaging effects. The Poorvu Center at Yale has compiled the following list:

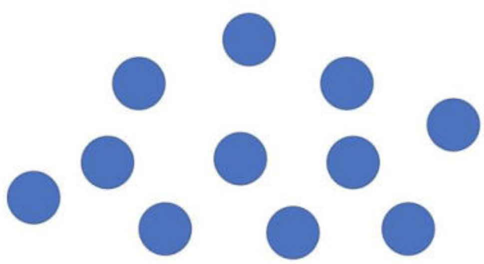
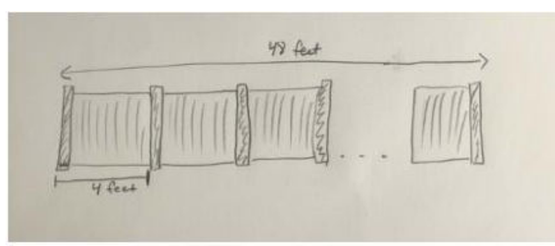
Use a Rubric or Table of Specifications	Instructors can use a rubric to lay out expected performance criteria for a range of grades. Rubrics will describe what an ideal assignment looks like, and "summarize" expected performance at the beginning of term, providing students with a trajectory and sense of completion.
Design Clear, Effective Questions	If designing essay questions, instructors can ensure that questions meet criteria while allowing students freedom to express their knowledge creatively and in ways that honor how they digested, constructed, or mastered meaning.
Assess Comprehensiveness	Effective summative assessments provide an opportunity for students to consider the totality of a course's content, making broad connections, demonstrating synthesized skills, and exploring deeper concepts that drive or found a course's ideas and content.
Make Parameters Clear	When approaching a final assessment, instructors can ensure that parameters are well defined (length of assessment, depth of response, time and date, grading standards); knowledge assessed relates clearly to content covered in course; and students with disabilities are provided required space and support.
Consider Blind Grading	Instructors may wish to know whose work they grade, in

	order to provide feedback that speaks to a student's term-long trajectory. If instructors wish to provide truly unbiased summative assessment, they can also consider blind grading.

One of the problems with a classroom approach based upon frequent grading is that teachers are using summative measures hoping they will have a formative effect and impact learning. One alternative to this approach is standards-based grading, which can be used in ways that support formative and summative assessment.

Examples of summative questions, from primary, upper elementary, middle school and high school, are given below.

Summative assessment questions:

Primary:	**Upper elementary:**
You have a collection of objects and your friend gives you 6 more. How many do you have and how do you know? Explain your reasoning using words, pictures and numbers. 	You have a 48 foot-long fence made up of 4 foot panels. How many 4 foot panels are there? How do you know? Write a number sentence showing the calculation needed for this question. Fully explain how your number sentence models this situation. 

Middle School:	**High School:**
A point is located at -17 on a number line. If you add 8 to -17 and move the point, where will it be located? Draw the number line showing the movement and write a number sentence that represents the movement of the point. What whole number is between? Make a convincing argument proving how you know. Explain your reasoning fully.	F(x) = 3x+2 where the domain is over the interval [0,7]. Graph the function and include a table of values showing the integer ordered pairs. Write a story that might be modeled by this function. Explain how your story models the function.

The MARS collaborative also provides conceptually oriented summative assessments, for example:

Middle School:

- Candle box: https://www.map.mathshell.org/tasks.php?unit=ME01&collection=9
- Hot under the collar: https://www.map.mathshell.org/tasks.php?unit=ME04&collection=9

High School:

- Propane tanks: https://www.map.mathshell.org/tasks.php?unit=HE16&collection=9
- Skeleton tower: https://www.map.mathshell.org/download.php?fileid=810

More examples are provided here: https://www.map.mathshell.org/tasks.php

Re-taking Assignments and Tests

Even when assignments and tests are used frequently, they can still provide a valuable learning experience for students when they are not seen as the end to a learning cycle. Some teachers believe that retaking work is unfair, as students may go away and learn on their own what they need to improve their grade but such efforts, are, at their core, about learning, and should be valued. Some teachers believe that if learners can retake and get full marks on their second attempt, it encourages students to take initial assessments less seriously, but this is not how students approach such opportunities. Allowing students to retake work sends an important growth mindset message, and encourages further learning. Just as career mathematicians are constantly revising their work and conjectures, students should be allowed the same fluidity in their own learning process. See the snapshot below for an example of how retaking a test can enable further learning.

> Allowing students to resubmit any work or test is the ultimate growth mindset message, focusing assessment upon *learning*, rather than performance.

Snapshot

Kaj has noticed that, for some of her students, the unit tests are anxiety-inducing—both in the taking of the tests and in receiving potentially low scores a few days later. In talking with an English language arts/literacy (ELA) peer teacher, the subject of testing came up, and her peer pointed out that drafts and revisions are the norm in ELA. Kaj wondered if embedding a revision cycle into the testing could help her students with test anxiety and with long-term retention. For her next unit test, she announces to the class that they will have the opportunity to revise their work on any items that they lost significant points. In the week before the test, she overheard some of her students mentioning that they might just “wing it” since they can just retake items later. She decided that a few rules were in order: when taking the test, an attempt must be made and answer found on all problems; a revision included three components: a correct solution with all steps shown, an annotated version of the original work with explanation of what was overlooked or missed, and a citation of the resource used—such as page number or class notes. On testing day, she noticed that for those students that typically struggled, they seemed to be writing more and leaving fewer questions unanswered. During grading she was careful to give written feedback (see earlier Diagnostic Comments section) that was both positive and constructive so students were more inclined to revise their work rather than scrapping it entirely, if possible. As the revisions came in, Kaj was heartened to see that her students improved upon their work considerably, and their scores reflected this improvement. She also noticed that, for many of her students, the revision process enabled better retention in the long term. As Kaj made further changes to the system, including a limit to how many problems could be revised, as well as instituting a peer checking system, she was able to address the extra grading time for herself as well as some of the complaints about fairness she overheard from a few parents. For the next year, she planned on including good study practices in the lead-up to a test, and having her students talk with a classmate to help identify which topics were most difficult for them. Overall, she felt that developing these types of reflection, self-awareness and anticipation skills in her students will bode well for them with future learning experiences.

Portfolios

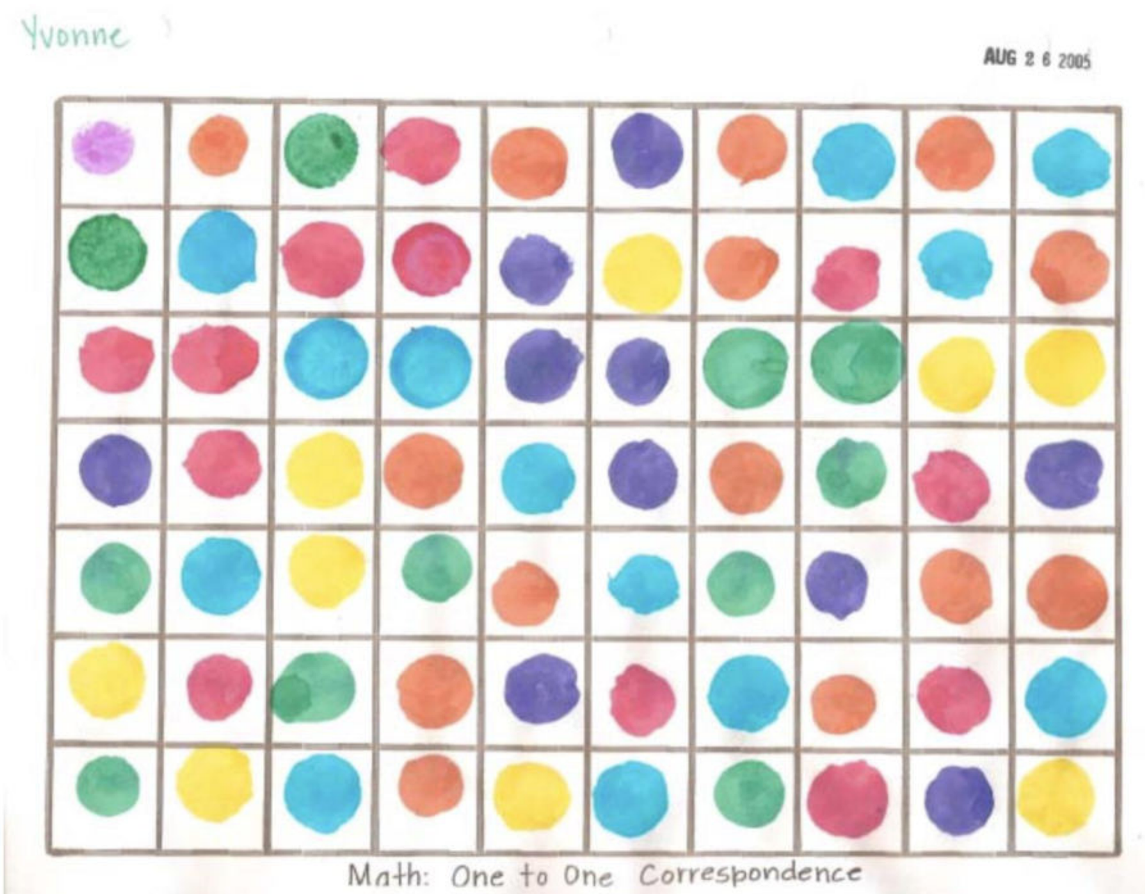
Perhaps the most comprehensive way to assess student learning is through a portfolio—a collection of work that communicates students' activities over a length of time. It could include project work, photographs, audio samples, letters, digital artifacts and other records of mathematical work. Portfolios allow students to choose and assemble their best work, selecting the contents and reflecting on the reasons for their inclusion. Portfolios are particularly appropriate ways of assessing data science projects. Students should have the option of demonstrating their math knowledge of math concepts through the use of their home language.

Advice and examples of portfolios are given here:
https://www.geneseo.edu/sites/default/files/sites/education/p12resources-portfolio-assessment.pdf

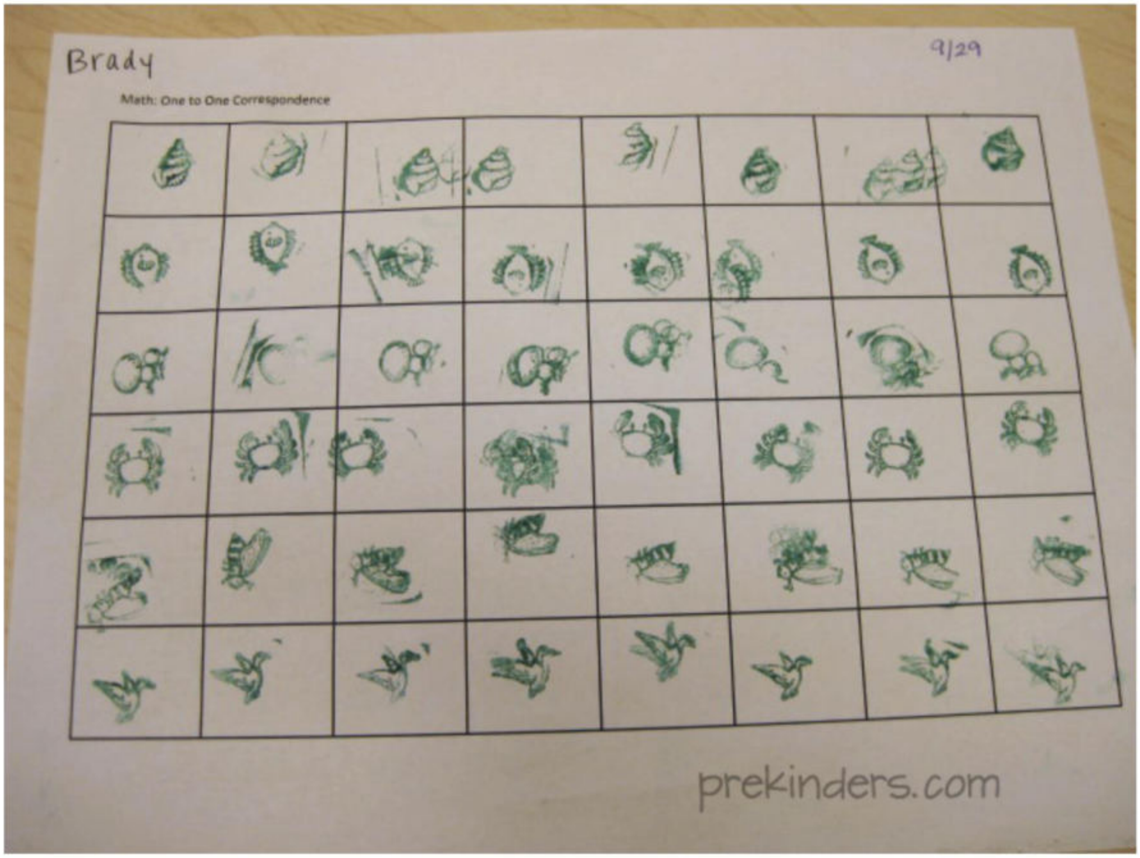
Advice on web tools for student portfolios is provided here:
https://www.edutopia.org/blog/web-tools-for-student-portfolios-dave-guymon

Portfolios can be scored using well-developed rubrics or criteria. They can have good value when used as a way of communicating student progress to parents. Ideally, they tell a story of student growth in learning the content and practices of mathematics. The detail can help parents support their students' learning and expand collaboration between schools and families. In distance learning settings, portfolios can provide a powerful means for students to demonstrate understanding and knowledge, and can be easily compiled with the use of technology.

Examples of Pre-K Mathematics Portfolios: https://www.prekinders.com/math-portfolios/ and the following are examples of tasks a kindergarten teacher included in her student portfolio:

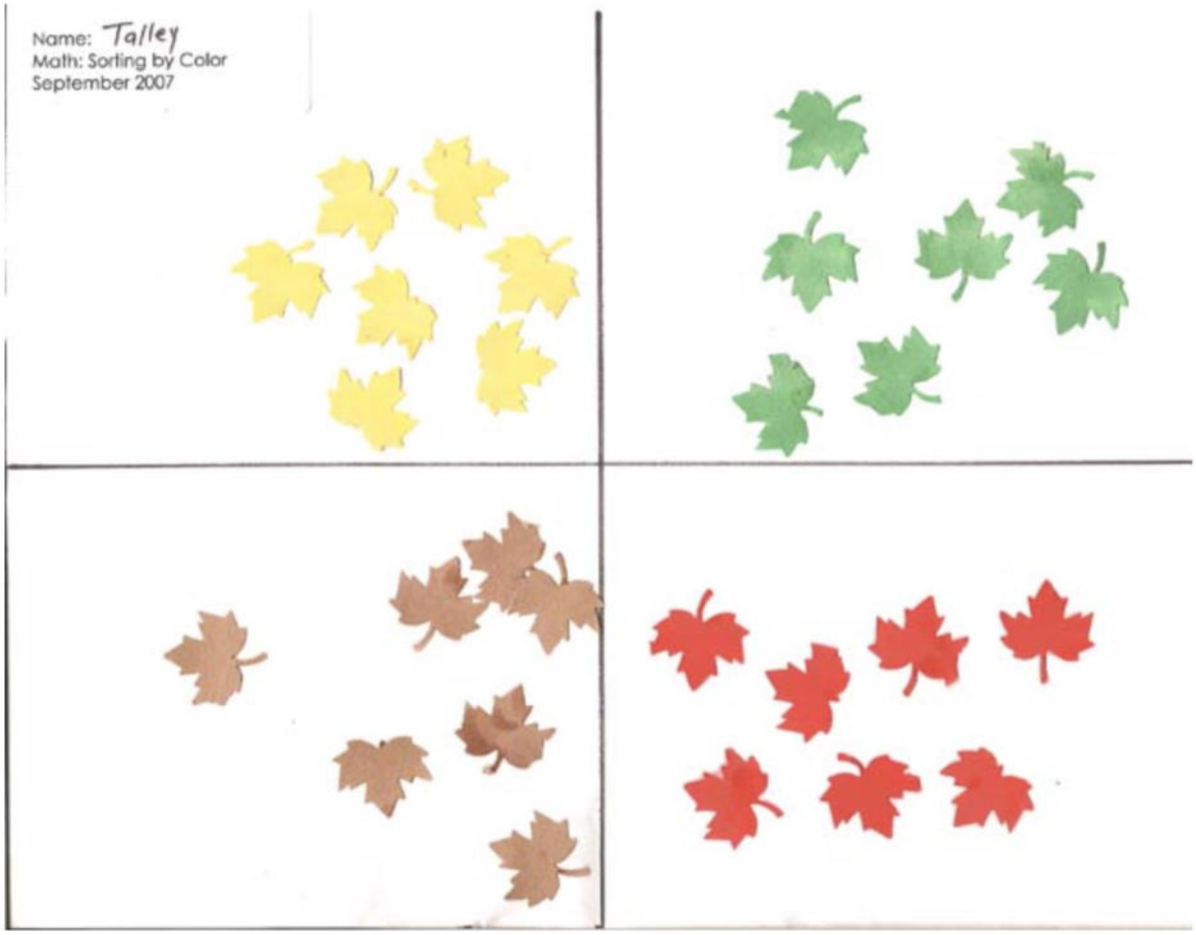


One to One Correspondence: stamp bingo dot markers in squares



One to one correspondence with rubber stamps

Represent numbers with drawing



Sorting paper cutouts by color

Examples of middle and high school portfolios including social justice examples can be seen here: https://sites.google.com/a/eschs.net/www/mathportfolios@eastside and advice for high school teachers designing portfolio assessments is given here: http://jfmueller.faculty.noctrl.edu/toolbox/examples/seaver/geometryportfolio.htm

Smarter Balanced Assessment Consortium and the CAASPP

California's participation in the Smarter Balanced Assessment Consortium has resulted in a statewide assessment program, known as the California Assessment of Student Performance and Progress (CAASPP). The CAASPP is designed to measure students' and schools' progress toward meeting the goals of the *California Common Core State Standards for Mathematics* (*CA CCSSM*) for grades three through eight and in grade eleven. Information on the CAASPP is available at http://www.cde.ca.gov/ta/tg/ca/.

Smarter Balanced assessments, and specifically the CAASPP, require students to think critically, solve problems, and show a greater depth of knowledge.

In measuring students' and schools' progress toward meeting the *CA CCSSM*, there are three key aspects of the CAASPP: computer-based, computer-adaptive and varied items.

- *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 adjust 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. Sample performance tasks can be found at https://understandingproficiency.wested.org/. Finally, the assessments feature technology-enhanced items that aim to provide evidence of mathematical practices. 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 uses the cluster headings in the *CA CCSSM* as the targets of assessment for generating evidence for the claim. 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*: SMP.5, SMP.6, SMP.7, and SMP.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. Many resources for modeling tasks exist, including this from Illustrative Math https://im.kendallhunt.com/HS/teachers/1/prompts.html

.

Interim Assessment

Interim assessments allow teachers to check students' progress at mastering specific concepts at strategic points throughout the year. Teachers can use this information to support their instruction and help students meet the challenge of college- and career-ready standards. Smarter Balanced offers the following interim assessments:

Interim Comprehensive Assessments (ICAs) that test the same content and report scores on the same scale as the summative assessments. Interim Assessment Blocks (IABs) that focus on smaller sets of related concepts and provide more detailed information for instructional purposes. Focused IABs that assess no more than three assessment targets to provide educators with a finer grained understanding of student learning. See https://contentexplorer.smarterbalanced.org/test-development for more information on interim assessments.

Conclusion

Assessment in mathematics is in a period of transition, from tests of fact-based skills to multi-faceted measures of sense-making, reasoning and problem-solving. In this way, there is a growing alignment between how mathematics is being taught, and how it is being tested. A comprehensive system of assessment should provide all stakeholders with the levels of detail to make informed decisions. Educators, administrators, and policymakers should focus on the most efficient use of assessment in mathematics learning, one that maximizes the amount of learning each child is capable of, while minimizing the socio-cultural effects of repeated testing. At the most fundamental level, each stakeholder has an important role in supporting classroom teachers' use of assessment in making the critical minute-by-minute decisions that afford better learning for all students in their care. All stakeholders working collaboratively within a system of assessment will ensure that all students in California have access to the rich mathematical ideas and practices set forth in the *CA CCSSM*.

References

Aguirre, J., Mayfield-Ingram, K., & Martin, D. (2013). *The impact of identity in K-8 mathematics: Rethinking equity-based practices*. The National Council of Teachers of Mathematics.

Banks, J. (2014). Series Forward. In N. Nasir, C. Cabana, B. Shreve, E. Woodbury, & N. Louie (Ed.) *Mathematics for equity: A framework for successful practice*. New York: Teachers College Press.

Black, P. J., & Wiliam, D. (1998, October). Inside the black box: Raising standards through classroom assessment. Phi Delta Kappan, 139–148.

Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2002). Working inside the black box: Assessment for learning in the classroom. London: Department of Education & Professional Studies, King's College.

Boaler, J. (2014). Research suggests that timed tests cause math anxiety. *Teaching children mathematics*, *20*(8), 469–474.

Boaler, J. & Staples, M (2014.) Creating Mathematical Futures through an Equitable Teaching Approach. *Mathematics for equity: A framework for successful practice*, 129–144.

Boaler, J. (2003). When learning no longer matters: Standardized testing and the creation of inequality. *Phi Delta Kappan*, *84*(7), 502–506.

Brookhart, S. M., Guskey, T. R., Bowers, A. J., McMillan, J. H., Smith, J. K., Smith, L. F., Stevens, M. T. & Welsh, M. E. (2016). A century of grading research: Meaning and value in the most common educational measure. *Review of Educational Research*, *86*(4), 803–848. https://www.jstor.org/stable/44668237

Cabana, C., Shreve, B., Woodbury, E., & Louie, N. (2014). *Mathematics for equity: A framework for successful practice*. Teachers College Press.

Carpenter, T. P., Fennema, E., Franke, M. L., Levi, L., & Empson, S. B. (1999). Children's mathematics. *Cognitively Guided*.

Cohen, E. G., & Lotan, R. A. (2014). *Designing groupwork: Strategies for the heterogeneous classroom Third Edition*. Teachers College Press.

DeSilva, E. (2020). Students in the Center. In Gardner, T. (Ed). *Leading in the Belly of the Beast: School Leadership in a School System Designed to Fail*. (pp. 95–112). Lanham Maryland: Rowman & Littlefield.

Feldman, J. (2019). Grading for Equity, What It Is, Why it Matters, and How it Can Transform Schools and Classrooms. Thousand Oaks, CA: Corwin.

Iamarino, D. L. (2014). The benefits of standards-based grading: A critical evaluation of modern grading practices. *Current Issues in Education*, *17*(2).

Kling, G., & Bay-Williams, J. M. (2014). Assessing basic fact fluency. *Teaching children mathematics*, *20*(8), 488–497.

Lambert, R. (2020). Increasing access to Universally Designed mathematics classrooms. Stanford, CA: PACE. https://edpolicyinca.org/publications/increasing-access-universally-designed-mathematics-classrooms

Leinwand, S., Huinker, D., & Brahier, D. (2014). Principles to actions: Mathematics programs as the core for student learning. *Mathematics teaching in the Middle school*, *19*(9), 516–519. doi:10.5951/mathteacmiddscho.19.9.0516

Linquanti, R. 2014. Supporting Formative Assessment for Deeper Learning: A Primer for Policymakers.

Meyer, A., Rose, D. H., & Gordon, D. T. (2014). *Universal design for learning: Theory and practice*. CAST Professional Publishing. http://udltheorypractice.cast.org/

Muñoz, M. A., & Guskey, T. R. (2015). Standards-based grading and reporting will improve education. *Phi Delta Kappan*, *96*(7), 64–68. https://doi.org/10.1177/0031721715579043

Schoen, H. L., Cebulla, K. J., Finn, K. F., & Fi, C. (2003). Teacher variables that relate to student achievement when using a standards-based curriculum. *Journal for Research in Mathematics Education*, 228-259. https://www.jstor.org/stable/30034779

Selbach-Allen, M. E., Greenwald, S. J., Ksir, A. E., & Thomley, J. E. (2020). Raising the Bar with Standards-Based Grading. *PRIMUS*, 1–17. https://doi.org/10.1080/10511970.2019.1695237

Townsley, M. & Buckmiller, T. (2016, January). What Does the Research Say about Standards-Based Grading? [Blog post] Retrieved from http://mctownsley.net/standards-based-grading-research/

Tsu, R., Lotan, R., & Cossey, R. (2014). Building a vision for equitable learning. *Mathematics for equity: A framework for successful practice*, 129–144.

Zwiers, J., Dieckmann, J., Rutherford-Quach, S., Daro, V., Skarin, R., Weiss, S., & Malamut, J. (2017). Principles for the Design of Mathematics Curricula: Promoting Language and Content Development. Retrieved from Stanford University, UL/SCALE website: http://ell.stanford.edu/content/mathematics-resources-additional-resources