# Chapter Eight

of the

# English Language Arts/ English Language Development Framework

*for California Public Schools*
*Kindergarten Through Grade Twelve*

**Adopted by the California State Board of Education, July 2014**

**Published by the California Department of Education**
**Sacramento, CA**

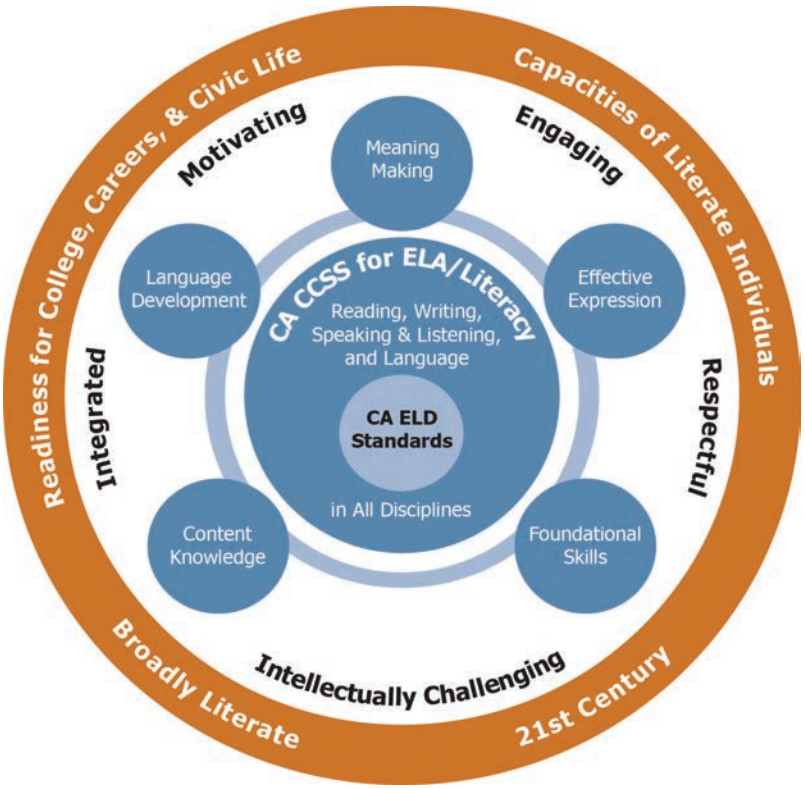
# Assessment


| Page | Chapter at a Glance |
|---|---|
| 822 | **Purposes of Assessment** |
| 825 | **Assessment Cycles** |
| 829 | Short-Cycle Formative Assessment |
| 835 | Medium-Cycle Assessment |
| 835 | End-of-Unit Assessments |
| 836 | Interim or Benchmark Assessments |
| 838 | Assessing ELD Using Medium-Cycle Evidence |
| 839 | Long-Cycle Assessment |
| 841 | Ensuring Accessibility for ELs on Long-Cycle Assessments |
| 842 | Additional Methods of Medium- and Long-Cycle Assessment |
| 842 | Rubrics |
| 844 | Portfolios |
| 845 | **Student Involvement** |
| 846 | Feedback |
| 847 | Teacher Feedback |
| 848 | Peer Feedback |
| 849 | Self-Assessment |
| 850 | **Assessment of ELD Progress** |
| 851 | Assessing ELD Progress in Writing |
| 856 | Assessing ELD Progress in Oral Language |
| 858 | **Assessment for Intervention** |
| 859 | Universal Screening (Medium Cycle) |
| 859 | Diagnostic Assessment (Medium Cycle) |
| 860 | Progress Monitoring (Short or Medium Cycle) |
| 861 | **Mandated California Assessments** |
| 864 | Computer Adaptive Tests |
| 864 | Performance Tasks |
| 865 | Assessments for Students with Significant Cognitive Disabilities |
| 865 | Biliteracy Assessment |
| 866 | English Language Proficiency Assessment |
| 867 | **Technical Quality of Assessments** |
| 868 | Elements of Technical Quality |

| Page | Chapter at a Glance (cont.) |
|---|---|
| 868 | Validity |
| 868 | Reliability |
| 869 | Freedom from Bias |
| 870 | Technical Quality and Formative Assessment |
| 872 | **Conclusion** |
| 873 | **Works Cited** |


Student achievement of the CA CCSS for ELA/Literacy and the CA ELD Standards depends on educators' skilled use of assessment information. With the institution of these standards, the landscape of assessment and accountability in California is experiencing a dramatic shift. Not only do the standards present new goals for California educators as depicted in the outer ring of figure 8.1, but the implementation of the California Assessment of Student Performance and Progress (CAASPP) system represents a major shift in the intent of statewide assessment: "It is the intent of the Legislature . . . to provide a system of assessments of pupils that has the primary purposes of assisting teachers, administrators, and pupils and their parents; improving teaching and learning; and promoting high-quality teaching and learning using a variety of assessment approaches and types" (*Education Code* 60602.5[a]). This shift is consonant with major emphases in California's standards for college and career readiness: a renewed focus on purposeful and deeper learning for students and their teachers, strong collaboration and partnerships at all levels of education, and a culture of continuous growth based on reflective practice.

***Figure 8.1. Circles of Implementation of ELA/Literacy and ELD Instruction***



This chapter describes the scope of assessment and its skilled use to support student achievement of the CA CCSS for ELA/Literacy and the CA ELD Standards—and ultimately the overarching goals of ELA/literacy and ELD instruction: students develop the readiness for college, careers, and civic life; attain the capacities of literate individuals; become broadly literate; and acquire the skills necessary for living and learning in the 21st century. (See outer ring of figure 8.1.) Both sets of standards, as discussed throughout this *ELA/ELD Framework*, constitute shifts that have significant implications for assessment.

*Both sets of standards, as discussed throughout this* ELA/ELD Framework, *constitute shifts that have significant implications for assessment.*

From the outset, the coherent structure of the CA CCSS for ELA/Literacy and CA ELD Standards from kindergarten through grade twelve lends itself to effective assessment practices. Mapping the CA CCSS for ELA/Literacy within each strand (Reading, Writing, Speaking and Listening[1], and Language) backwards from the College and Career Readiness (CCR) Anchor Standards, makes clear what students are to know and be able to do at each grade and also demonstrates the relatively small number of broad competencies to assess as students move from novice to expert. Similarly, the organization of the three parts of the CA ELD Standards ("Interacting in Meaningful Ways," "Learning About How English Works," "Using Foundational Literacy Skills") helps teachers make important instruction and assessment decisions for ELs by grade level *and* English language proficiency level.

Meaningfully, both sets of standards encompass the full spectrum of language and literacy competencies from kindergarten through grade twelve, thereby providing many opportunities for students to apply and transfer skills from the earliest grades. The standards encourage educators to think broadly and plan instruction comprehensively. "[E]ach standard need not be a separate focus for instruction and assessment. Often, several standards can be addressed by a single, rich task [, so that] students can develop mutually reinforcing skills and exhibit mastery . . . across a range of texts [and tasks]" (CDE 2013a, 4–5).

*Importantly, the standards recommend that language and literacy learning be connected with the academic disciplines from the earliest grades onward. Assessment, then, should enable educators to determine a student's trajectory in developing proficiency in language and literacy within and across the years and the disciplines.*

Importantly, the standards recommend that language and literacy learning be connected with the academic disciplines from the earliest grades onward. Assessment, then, should enable educators to determine a student's trajectory in developing proficiency in language and literacy within and across the years and the disciplines.

The chapter begins with an explanation of the different purposes of assessment—both *for* and *of* learning. Cycles of assessment—short, medium, and long—are then discussed, including the types and purposes of assessment within each time frame and the decisions that each assessment type can inform. Snapshots of teacher use of assessment are included throughout the discussion of assessment cycles. The role of student involvement and feedback in assessment is highlighted, followed by guidance for assessment of ELD progress and descriptions of assessment for intervention. In addition, the chapter

1 As noted throughout this framework, speaking and listening should be broadly interpreted. Speaking and listening should include deaf and hard of hearing students using American Sign Language (ASL) as they primary language. Students who are deaf and hard of hearing who do not use ASL as their primary language but use amplification, residual hearing, listening and spoken language, cued speech and sign supported speech, access the general education curriculum with varying modes of communication.

presents a brief overview of mandated statewide assessments and concludes with a consideration of the technical quality of assessments to ensure that assessments yield accurate information for their intended purposes.

This chapter can be used in several ways. As a source of professional learning for teachers and school and district leaders, the chapter plays a critical role in strengthening educators' assessment literacy—their knowledge and understanding of assessment practices and appropriate uses of assessment evidence to shape powerful instruction. The chapter provides teachers and leaders a structure for examining the types of assessment practices and sources of assessment evidence currently in use in schools and for proposing needed additions and adjustments. This chapter features formative assessment as a process and recommends that it be the focus of in depth professional learning and support, including dialogue with peers, classroom practice of new approaches, and coaching.

# Purposes of Assessment

Assessments are designed and used for different purposes. For example, an annual assessment designed to assess how well students have met a specific standard (e.g., CA CCSS for ELA/Literacy RI.4.8: *Explain how an author uses reasons and evidence to support particular points in a text*) does just that: It indicates whether students have met a specific standard. However, this assessment does not diagnose a particular reading difficulty a fourth-grade student is experiencing in achieving the standard. Nor does it provide substantive insights into how a student is beginning to understand what constitutes evidence in a specific text. In the use of any assessment, a central question is, "Is this assessment being used for the purpose for which it is intended?"

*Assessment has two fundamental purposes: One is to provide information about student learning minute-by-minute, day-to-day, and week-to-week so that teachers continuously adapt instruction to meet students' specific needs and secure progress.*

Assessment has two fundamental purposes: One is to provide information about student learning minute-by-minute, day-to-day, and week-to-week so that teachers continuously adapt instruction to meet students' specific needs and secure progress. This type of assessment is intended to assist learning and is often referred to as formative assessment or assessment ***for*** learning. Formative assessment occurs in real time—during instruction—while student learning is underway (Allal 2010; Black and Wiliam 1998; Bell and Cowie 2000; Heritage 2010; Shepard 2000, 2005b). For example, a third-grade teacher working with small groups of students on distinguishing their point of view from a particular author's viewpoint gains insights into students' developing skills through the use of strategic questions and uses students' responses to adjust instruction.

Although discussed further in the next section, formative assessment is briefly defined in figure 8.2.

***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.

A second purpose of assessment is to provide information on students' current levels of achievement after a period of learning has occurred. Such assessments—which may be classroom-based, districtwide, or statewide—serve a summative purpose and are sometimes referred to as assessments ***of*** learning. They 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). Inferences made by teachers from the results of these assessments are used to make decisions about student placement, instruction, curricula, interventions, and to assign grades. For example, the current state assessment of English language proficiency, the California English Language Development Test (CELDT), measures an EL's annual progress in attaining proficiency. School districts use the results of the annual assessment to make decisions about the ongoing instructional placement or possible reclassification of ELs. The English Language Proficiency Assessments for California (ELPAC) are scheduled replace the CELDT in 2017 or 2018. (See the section on English language proficiency assessments in this chapter.)

*A second purpose of assessment is to provide information on students' current levels of achievement after a period of learning has occurred. Such assessments—which may be classroom-based, districtwide, or statewide—serve a summative purpose and are sometimes referred to as assessments* **of** *learning.*

As part of a balanced and comprehensive assessment system, assessment ***for*** learning and assessment ***of*** learning are both important. While assessment(s) of learning usually involve a tool or event *after* a period of learning, assessment for learning is a process. Evidence-gathering strategies that are truly formative yield information that is *timely* and *speci ic* enough to assist learning as it occurs. Figure 8.3 presents the key dimensions of assessment *for* and *of* learning and highlights their differences.

***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.

As figure 8.3 illustrates, assessment for learning—comprising key practices of the formative assessment process—occurs during instruction (or while learning is happening) and addresses students' immediate learning needs. Intertwined and inseparable from teachers' pedagogical practice, formative assessment is a high priority. It is especially important as teachers assess and guide their students to develop and apply a broad range of language and literacy skills. The special note (*) in figure 8.3 indicates that some assessments of learning can be used for formative purposes. In other words, they can be used to inform future teaching and learning and not simply to report on achievement or progress. This is only the case *if* the evidence-gathering tool addresses *both* the focus of instruction of the previous unit *and* immediate future learning goals.

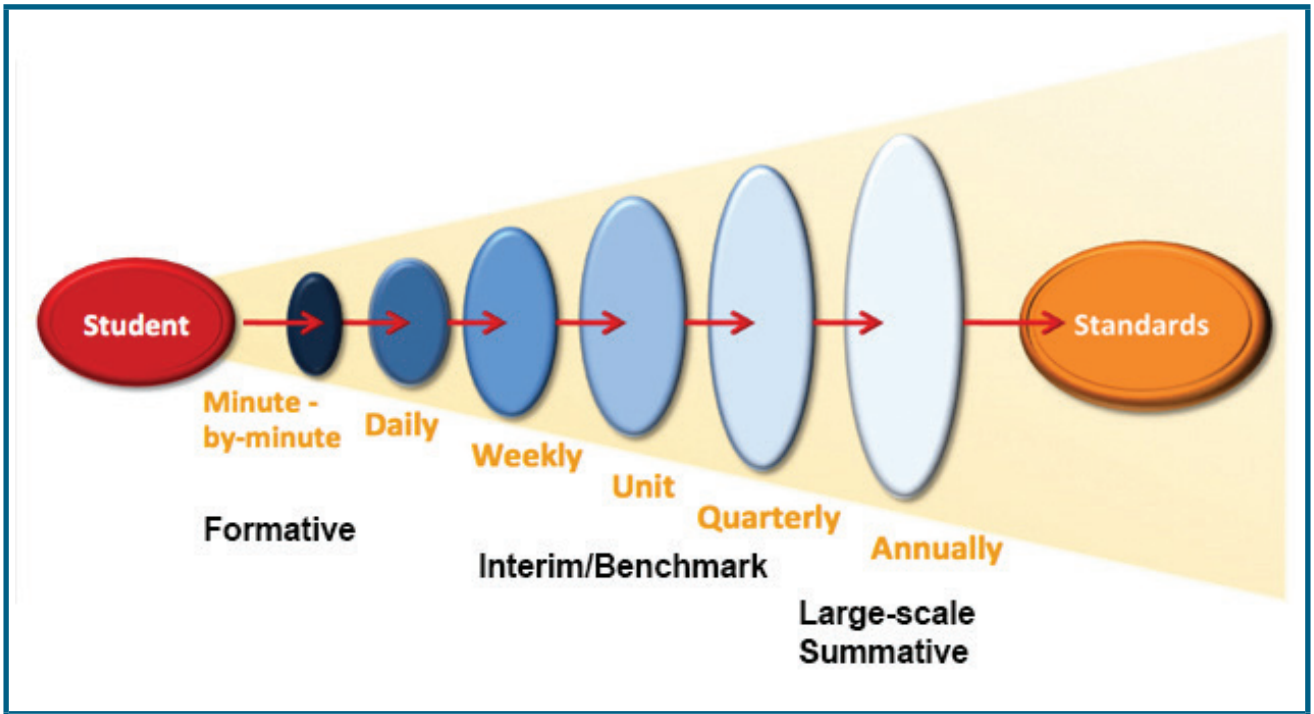
*Intertwined and inseparable from teachers' pedagogical practice, formative assessment is a high priority. It is especially important as teachers assess and guide their students to develop and apply a broad range of language and literacy skills.*

School leaders and professional learning providers consider the support that educators require to understand and implement the formative assessment process fully, as well as to use interim/benchmark and summative assessments effectively. Importantly, educators (classroom teachers, specialists, administrators, and others) interpret assessment evidence in order to plan instruction and respond pedagogically to emerging student learning. Collaborative professional environments, such as communities of practice, are the nexus of learning, and the work teachers do relative to assessment evidence is part of an ongoing cycle of inquiry. (See chapter 11 in this *ELA/ELD Framework.*) To optimize instructional decision making relative to the CA CCSS for ELA/Literacy and the CA ELD Standards, teachers and leaders make full use of assessment for both formative and summative purposes.

# Assessment Cycles

One way to consider assessment for different purposes is to conceptualize assessment as operating in different cycles: short, medium, and long (Wiliam 2006). Figure 8.4 presents a range of assessments within a comprehensive assessment system. Those assessments that are more proximate to student learning (i.e., minute-by-minute, daily, weekly) operate in a short cycle because they address a short period of teaching and learning. Short-cycle assessment serves a formative purpose because its intended use is to inform immediate teaching and learning. Assessments administered at the end of the year operate in a long cycle because they cover a much longer period of learning. Long-cycle assessments are primarily used for summative purposes.

***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.

Occupying a middle position between short-cycle (formative) and annual (summative) assessment is interim/benchmark assessment: "assessments administered periodically throughout the school year, at specified times during a curriculum sequence to evaluate students' knowledge and skills relative to an explicit set of longer-term learning goals" (Herman, Osmundson, and Dietel 2010, 1). In figure 8.4, classroom summative assessments are referred to as unit assessments (although they could also occur in shorter time frames), and interim/benchmark assessments are referred to as quarterly assessments. Such periodic assessments operate in a medium cycle because they address longer-term goals than those assessments more proximate to student learning but not as long-term as annual assessments. Classroom summative or interim/benchmark assessments are generally used for summative purposes—evaluating what has been learned—although they may be used for formative purposes if they inform decisions that teachers and instructional leaders make within the school year regarding curricula, instructional programs and practices, and professional learning to improve future student learning. However, classroom summative or interim/benchmark assessments are distinct from the formative assessment process because, by their design and intended use, they do not inform immediate teaching and learning. Unit assessments primarily serve a summative function but can serve a formative purpose if the teacher can act on the assessment information to support improved learning in a future unit. Progress-monitoring assessments can be short, medium, or long cycle, depending on whether they are administered after a shorter or longer period of instruction, and serve both a formative and summative function. (For more information on screening, diagnostic assessment, and progress monitoring, see subsequent sections of this chapter).

Assessments within each cycle function best when they are part of a comprehensive, coherent, and continuous system of assessment that provides ongoing information to teachers throughout the year (NRC 2001). Within such systems, minute-by-minute, daily, and weekly assessment feeds into unit assessment, which, in turn, feeds into periodic (e.g., end-of-unit, quarterly interim/benchmark) assessments, and multiple interim assessments feed into the annual assessment of the standards. A comprehensive, coherent, and continuous system of assessment provides mutually complementary views of student learning, ensures that assessment within each cycle is focused on the same ultimate goal (achievement of standards), and pushes instruction and learning in a common direction (Herman 2010).

Each assessment cycle provides information at varying levels of detail, and inferences drawn from the assessment results are used to address specific questions about student learning and inform a range of decisions and actions. Figure 8.5 summarizes the types and purposes of the assessments within each assessment cycle.

***Figure 8.5. Types and Uses of Assessments Within Assessment Cycles***

| *Cycle* | *Methods* | *Information* | *Uses/Actions* |
|---|---|---|---|
| **Short** | | | |
| **Minute-by-Minute** | • Observation<br>• Questions (teachers and students)<br>• Instructional tasks<br>• Student discussions<br>• 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:<br>• Observation<br>• Questions (teachers and students)<br>• Instructional tasks<br>• Student discussions<br>• Written work/ representations<br>• 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<br>• Instructional adjustments in this or the next lesson<br>• Find out more<br>• Feedback to class or individual students (oral or written) |
| **Week** | • Student discussions and work products<br>• 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<br>• Feedback to students (oral or written) |

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

## Short-Cycle Formative Assessment

Short-cycle formative assessment is a process used by teachers and students *during instruction* that provides feedback to adjust ongoing teaching and learning to improve student achievement of intended instructional outcomes (McManus 2008). It occurs when evidence of learning is gathered minute-by-minute, daily, and weekly from a variety of sources during ongoing instruction for the purpose of moving learning forward to meet short-term goals (i.e., lesson goals) (Black and Wiliam 1998; Council of Chief State School Officers Formative Assessment State Collaborative 2006; Heritage 2010; Popham 2010). In the remainder of this chapter, this short-cycle formative assessment process is referred to as formative assessment.

*Short-cycle formative assessment is a process used by teachers and students* during instruction *that provides feedback to adjust ongoing teaching and learning to improve student achievement of intended instructional outcomes.*

This type of assessment provides the most detailed information for teachers and their students. The idea of formative assessment, or assessment *for* learning, does not apply to a specific tool or assessment. A tool or assessment can be used for formative assessment purposes, but only if it provides actionable information about students' learning relative to the desired lesson goal and teachers can use it immediately to adjust their instruction. Many assessments marketed under the formative assessment label do not provide information about students' learning needed to adjust instruction and guide students' learning as it occurrs (Perie, Marion, and Gong 2009; Shepard 2005a).

The sources of evidence available to teachers in short-cycle formative assessment are what students do, say, make, or write (Griffin 2007). For example, sources of evidence can be teacher-student interactions fueled by well-designed questions (Bailey and Heritage 2008; Black, and others 2003), structured peer-to-peer discussions that the teacher observes (Harlen 2007), dialogues that embed assessment into an activity already occurring in the classroom (Ruiz-Primo and Furtak 2004, 2006 2007), student work resulting from well-designed tasks (Poppers 2011), and Web-based reading assessments that provide immediate feedback (Cohen, and others 2011).

The report of the Formative Assessment for Students and Teachers/State Collaborative on Assessment and Student Standards Project of the Council of Chief State School Officers emphasizes several features of formative assessment. First, "formative assessment is a *process* rather than a particular kind of assessment . . . There is no such thing as a 'formative test'" (McManus 2008, 3). Second, "the formative assessment process involves both teachers *and* students . . . , both of whom must be actively involved in the process of improving learning" (3). Third, teachers are clear about the ultimate goal of a unit and the sub-goals or stepping stones that are important along the way: ". . . from a learning progression teachers have the big picture of what students need to learn, as well as sufficient detail for planning instruction to meet short-term goals" (4). Fourth, success criteria and evidence of learning are laid out at the beginning of the project and reviewed along the way: ". . . teachers must provide the criteria by which learning will be assessed . . . using language readily understood by students, with realistic examples of what meets and does not meet the criteria."

*The sources of evidence available to teachers in short-cycle formative assessment are what students do, say, make, or write.*

Whatever the source of evidence, teachers construct or devise ways to elicit responses from students that reveal where they are in their learning and to use the evidence to move learning forward (Sadler 1989). Teachers are clear about the short-term learning goals (e.g., for a lesson) that cumulatively lead to students' attainment of one or more standards. They are also clear about the success criteria for the lesson goal—how students show they have met, or are on the way to meeting, the lesson goal. Teacheres then align the evidence-gathering strategy to the success criteria.

*Teachers are clear about the short-term learning goals (e.g., for a lesson) that cumulatively lead to students' attainment of one or more standards. They are also clear about the success criteria for the lesson goal . . .*

Questions that formative assessment can answer include the following:

- Where are my students in relation to learning goals for this lesson?
- What is the gap[2] between students' current learning and the goal?
- What individual difficulties are my students having?
- Are there any missing building blocks in their learning?
- What do I need to adjust in my teaching to ensure that students learn?

Information from formative assessment is used to make instructional adjustments in real time: to continue with the planned lesson or to provide feedback to students that helps them take steps to advance their learning. (Feedback to students is discussed in the student involvement section of this chapter.)

Importantly, teachers' inferences from formative assessment evidence and their resulting actions focus on individual students. The implication is not that instruction is necessarily provided on a one-to-one basis, but rather that individual needs are addressed in the context of a class of students. This orientation to individuals is necessary for students to have the opportunity to learn and progress equally (Heritage 2013). Accordingly, instruction is contingent on each student's current learning status. In other words, instruction is matched to where the students are so that they are assisted to progress and meet desired goals.

*Importantly, teachers' inferences from formative assessment evidence and their resulting actions focus on individual students.*

While formative assessment evidence is not aggregated in the same way as medium- and long-cycle assessment information, teachers can categorize individual student responses to look for patterns across the class or for particular students who are outliers. For example, after students have responded to a question about a text, a teacher can quickly categorize responses into those that demonstrate understanding, those that demonstrate partial understanding, and those that do not demonstrate understanding. The next day's instruction is then planned accordingly.

Teachers of ELs should take great care when making these formative assessment decisions. Depending on their level of English language proficiency, some ELs may not be able to fully express their ideas orally about a topic during a class discussion; however, this does not necessarily mean that they do not understand the topic. In addition, an informal observation indicating that ELs are not orally proficient in English should not determine how the students are taught reading in English.

2 The gap refers to the distance between where the students' learning currently stands at particular points in the lesson (a lesson can be several periods or days long) and the intended learning goal for the lesson. The purpose of short-cycle formative assessment is to close this gap so that all students meet the goal (Sadler 1989). This should not be confused with the term *achievement gap*, which refers to differences in summative educational outcomes among different subgroups of students.

English learners do not have to be proficient in oral English before they can learn to read in English (Bunch, Kibler, and Pimental 2012). Teachers use a combination of observations (e.g., during collaborative conversations among students about texts read) and informal inventories of reading (e.g., listening to students read aloud during small reading group time, asking specific comprehension questions to elicit student understandings) to determine how best to instruct their ELs and provide *just-in-time* scaffolding in reading. Furthermore, the CA ELD Standards indicate that all ELs, regardless of their level of English language proficiency, are capable of engaging in intellectually rich tasks at the same cognitive level as their English-proficient peers. With this aim, teachers use in-the-moment formative assessment practices to determine appropriate levels of scaffolding for ELs. (For more information on scaffolding, see chapter 2 in this *ELA/ELD Framework*.)

*Teachers of ELs should take great care when making these formative assessment decisions. Depending on their level of English language proficiency, some ELs may not be able to fully express their ideas orally about a topic during a class discussion; however, this does not necessarily mean that they do not understand the topic.*

Using the formative assessment process in an EL student's primary language, in contexts where teaching and learning use this resource (e.g., in an alternative program), may also offer instructionally actionable information. For example, newcomer ELs at the Emerging level of English language proficiency (e.g., students who have been in the U.S. for less than a year) may find it difficult to respond (in writing or orally) to a question about a science or history topic in English with the same level of detail as they are able to do in their primary language. Teachers can ask their newcomer EL students to quickly write responses to text-based questions first in their primary language (if they are literate) before they respond in written English. The two pieces of writing are then compared to identify similarities and differences between content knowledge and literacy in the primary language and English.

This technique is applied strategically so that teachers understand clearly what students know about particular topics and how well they are able to express their knowledge in English. Teachers also use this type of evidence to explicitly draw their EL students' attention to ways they can express through English writing or speaking what they already know and are able to convey in their primary language. While all teachers may not be able to provide this type of support themselves (e.g., when they are not proficient in students' primary languages), they can collaborate with other teachers, EL specialists, or community members to do so.

The use of technology that enables students to give immediate responses to teachers (e.g., clickers, mobile devices) helps teachers with large numbers of students gain an ongoing sense of students understanding during a lesson. For example, halfway through a lesson, a tenth-grade teacher asks three or four questions related to multiple-meaning words and word phrases in a literary text the class is analyzing. The results immediately appear as a pie chart on the Smart board. The teacher and students quickly see how the class responds and decide together if more work is needed in this area before the lesson progresses.

*The use of technology that enables students to give immediate responses to teachers (e.g., clickers, mobile devices) helps teachers with large numbers of students gain an ongoing sense of students understanding during a lesson.*

The following snapshots provide additional concrete examples of formative assessment in action.

## Snapshot 8.1. Formative Assessment in Grade Five

Fifth graders are working on the following CA CCSS for ELA/Literacy: (a) applying the reading standard for informational text: *explaining how an author uses reasons and evidence to support particular points in a text, identifying which reasons and evidence support which particular points* (RI.5.8); (b) the writing standard: *produce clear and coherent writing in which the development and organization are appropriate to task, purpose, and audience* (W.5.4); and (c) the language standard: *vocabulary use* (L.5.4-6), particularly transition words to help their writing flow logically. Students are writing an argument to encourage their readers to take more care of the natural environment. In their reading instruction, they analyzed a text to identify the location of *arguments*, *counterarguments*, and supporting *evidence*. In their writing, they are learning how to organize their arguments effectively.

While the students are involved in the independent writing part of the lesson, Ms. Hatwal sits with Bobby to discuss his writing progress. She has a ring binder open to a page with these headings at the top: *Child's Name/Date, Research Compliment, Teaching Point,* and *What's Next for this Child?* Further down the page is a self-adhesive note that lists five students' names, including Bobby's. She plans to meet with each of them during today's writing session.

Ms. Hatwal's initial purpose with Bobby is to follow up on feedback she provided him two days ago based on evidence she elicited from an interaction with him; in that interaction she determined that he needed to provide stronger sources of evidence to support his argument. On this occasion, she wants to see how he has used her prior feedback:

Ms. Hatwal: You're working on evidence? Tell me about it.

Bobby: I found good information in the book of the Environmental Protection Agency and on the Internet.

Ms. Hatwal: And what do you think about what you found so far? Do you think that it supports your argument?

Bobby: I guess . . . .

At this point, Ms. Hatwal reminds Bobby that the purpose of the evidence is to support his argument. She explains the meaning of "supporting an argument" in a way that is understandable to a fifth grader, by telling him: *You have to prove it with what is in the text or the readers may not believe you*. She asks him to read his argument aloud. Having established that the focus of his argument is to "stop dumping in the ocean because all the beautiful animals we see are going to start vanishing,"

Ms. Hatwal: So, what evidence did you find to support that claim—that all the animals will die if we don't stop dumping? What evidence did you find that will help you to strengthen that argument, or prove it to your readers?

Ms. Hatwal then helps Bobby recognize which of the information he has located is from a reliable source and is effective in supporting his argument. Satisfied that Bobby can move forward on his own to incorporate his evidence, she then asks him to review the organization of his argument and to let her know where he will place the evidence. When Bobby does this, it is evident to Ms. Hatwal that he has some confusion about the overall structure and that his writing needs to be reorganized. This is a moment in the interaction when she targets a teaching point for him. She reviews the organization with him and writes the organizational elements on a self-adhesive note and includes specific instructional support, such as putting the evidence in order to help the flow or adding transitional sentences.

### Snapshot 8.1. Formative Assessment in Grade Five (cont.)

Throughout this interaction, Ms. Hatwal makes notes in her ring-binder file. Under *Research Compliment* she writes that Bobby recognizes the reliability of his source. In the section labeled *Teaching Point* she writes that she explained how evidence supported his argument. Under the heading *What's Next for this Child?* she writes "organization and transitional sentences," noting that Bobby has problems organizing his writing to effectively convey his argument to the reader. By gathering evidence in the course of this interaction, Ms. Hatwal is able to match her teaching points to the individual student's needs. Additionally, after several interactions of this kind, she finds that there are common needs among several students and decides to pull them together for a mini-lesson.

### Snapshot 8.2. Formative Assessment in Grade Two

In a second-grade classroom that includes native English speaking children and children who are ELs, the children have been working on retelling folktales they have read together in class to convey the central message of the tale (RL.2.2). The EL children, in particular, have been working on using the past tense to indicate that the tales happened in the past (ELD. PII.2.3). In this lesson students are engaged in small group work, and during this time the teacher, Mr. Elfert, selects groups of three students to recount one of the folktales the class has read that week. In this situation, he wants to give each student sustained opportunities to use language while he and the others in the group listen. He asks the first student to begin, then after a while asks the second child to carry on and so forth. When the students have completed the retelling, Mr. Elfert asks them to say what they think the main message of the story is. Each child offers an opinion and a discussion follows about whether there is agreement on the main message. From the activity, Mr. Elfert has evidence that one student uses the past tense consistently and mostly with accuracy, while the other two do not. Two of the children are able to convey the message of the text, but another has not grasped it. After his discussion with the group, he makes quick notes about each student and briefly records his thoughts about subsequent instruction. He repeats this process with one additional group before the small group work time is over, and he plans more opportunities during the week to assess other small groups in the same way.

## Snapshot 8.3. Formative Assessment with Secondary EL Newcomers

In a secondary designated English Language Development (ELD) class, with newcomers whose experience in the U.S. ranges from three months to one year, Mrs. Rogers-Tsai works collaboratively with the science teacher, Miss Goodwin, to create a five-week unit on animal behavior with the purpose of guiding students through a deep exploration of the content through the language resources used to convey meaning. The two teachers have agreed that during science instruction, Miss Goodwin will provide appropriate and strategic support so EL students can fully participate in the science activities, gain understanding from the science textbook, and engage in collaborative discussions about the text and content. This strategic support includes using graphic organizers, providing increased opportunities for the students to discuss their ideas in small groups or pairs, and primary language support, including drawing attention to cognates and using texts in students' primary languages.

Mrs. Rogers-Tsai has agreed to analyze the science textbook and the activities the science teacher has designed in order to identify the language demands they present and then address the language demands more intensively during designated ELD instruction. This is the third class of the first week on the unit. Having formulated questions they would like to explore around the science topic, students perused a variety of texts on the topic to identify meanings and charted language (including phrasing and general academic and domain-specific vocabulary) they think is critical for conveying their understandings of the topic. They now work in pairs to collaboratively write a description about what they have learned so far about one aspect of animal behavior, using as much of the language they have charted as they can. The pairs write their description drafts on large sheets of paper, which they read to the class. Their peers are invited to ask questions and make comments. When one pair shares their description about animals and language, an animated conversation develops on whether animals have language. Julio explains the thinking that went into the description that caused the lively discussion.

Julio: First of all, I think that language is a way to *inform* others around you, your feelings or just a simple thing that you want to let know people what is the deal. And it can be *expressed* by saying it, watching a picture, or hearing it, you know what I'm saying? I don't know if you have heard about the kangaroo rat that stamps its feet to *communicate* with other rats. It's really funny cause we humans have more *characteristics* to *communicate* to each other, but we still have problems to understand other people. Characteristics like sound, grammar, pitch, and body language are some of them, while the rat only uses the foot (he stamps the ground).

Mrs. Rogers-Tsai, who has been recording in her notebook the language students use in the conversation, notes that Julio is using some of the academic language from the class chart in both his writing and speaking and has, more importantly, done an effective job of conveying his understanding of the information from his research and persuading his peers with evidence. Mrs. Rogers-Tsai decides to examine more closely the students' written descriptions, as well as the language they use in their conversations, in order to make decisions about what language features of the science texts to focus on as she progresses in the unit. She also plans to make a copy of her notes to share with Miss Goodwin when they meet later that week during collaboration time.

## Medium-Cycle Assessment

Assessments that teachers develop, or that are included in curricular materials and are administered at the end of a unit, are medium-cycle assessments. As noted previously, medium-cycle assessments occupy a middle ground between short-cycle formative assessment and long-cycle summative assessments. Some are used to inform instruction during the school year; others serve evaluative purposes.

### *End-of-Unit Assessments*

End-of-unit assessments serve a summative purpose to evaluate student achievement with respect to the goals of the unit. If such assessments are given to students before the end of the unit when there is still time to take instructional action before moving on to the next unit, then they also serve a formative purpose. In developing unit assessments, teachers ensure that the goals of the unit are clear and aligned to standards. In other words, what is to be assessed is well articulated and derived specifically from the standards and lesson planning. When teachers know *what* to assess, they can best determine *how* to assess. Teachers can then decide on the most effective way for students to demonstrate their achievement of the goals.

*In developing unit assessments, teachers ensure that the goals of the unit are clear and aligned to standards. In other words, what is to be assessed is well articulated and derived specifically from the standards and lesson planning.*

End-of-unit assessments help teachers answer questions such as the following:

- Have my students met the goals of the unit?
- Are there some students who need additional help to meet the goals of the unit?
- What help do they need?
- What improvements do I need to make in my teaching next time I teach this unit?

The following snapshot provides a concrete example of the use of end-of-unit (medium-cycle) assessment.

**Snapshot 8.4. End-of-Unit (Medium-Cycle) Assessment in Grade Seven**

In a seventh-grade classroom with native English speakers, recently reclassified ELs, and a group of ELs who are at the Expanding and Bridging levels of English language proficiency, Ms. Lambros has engaged students in a five week unit: Persuasion Across Time and Space: Analyzing and Producing Complex Texts. This unit addresses multiple CA CCSS for ELA/Literacy and CA ELD standards simultaneously and has four primary goals: (1) students read and analyze complex texts; (2) students identify and use evidence from informational texts in their written and oral work; (3) students participate in disciplinary practices highlighting language, purpose, and responsiveness to audience; and (4) students acquire history/social studies knowledge through content rich nonfiction.

**Snapshot 8.4. End-of-Unit (Medium-Cycle) Assessment in Grade Seven (cont.)**

During the course of the unit, with intentional and strategic scaffolding by Ms. Lambros and considerable involvement in collaborative groups, students engage in close reading, collaborative discussions, and analysis of the text organization, grammatical structures, and vocabulary of persuasive texts on relevant topics. In the final part of the unit, the students analyze the video, "The Girl Who Silenced the World for Five Minutes," compare and contrast persuasive techniques in the video to one of the texts they read, and produce a persuasive text of their own. The students' analyses of the video and written work serve as the summative assessment for the unit. Using the students' work, the teacher is able to make a determination about students' understanding of the purpose, organization, and structure of persuasive texts and their ability to use various language resources (including vocabulary, complex grammatical structures, connecting words and phrases) to write a coherent and cohesive persuasive piece for a public audience.

After reviewing students' responses, Ms. Lambros concludes that the students have made good progress toward meeting the goals of the unit, especially in regard to their understanding of persuasive techniques in different contexts (i.e., video and text). Examining her EL students' writing more closely, however, she notices that most of their writing is characterized by spoken, everyday language. In other words, their written arguments are not making use of connecting words and phrases (e.g., *for example*, *therefore*, *consequently*) to create cohesion, nor are they using many complex sentences to connect ideas and create relationships between them (e.g., *Even though governments are taking action*, it is not happening fast enough). This analysis of her students' writing informs Ms. Lambros's planning of subsequent lessons. She begins designing lessons in which she will show examples of cohesion and complex sentences that connect ideas, model how to *unpack* the meaning in the texts, collaboratively construct similar writing with the students, and provide students with guided practice in writing related to the unit topic. She also plans to draw her students' attention to various examples of persuasive language used in arguments and to observe how her students incorporated them into their own writing in upcoming units. In addition, she makes a note to address these linguistic features directly when she teaches the unit the following year.

**Resources**

Adapted from

Understanding Language. 2013. *Instructional Unit: Persuasion Across Time and Space.* Palo Alto, CA: Stanford University.

Video: *The Girl Who Silenced the World for 5 Minutes* (available at a number of sites, including www.youtube.com).

## *Interim or Benchmark Assessments*

Interim or benchmark assessments, such as the Smarter Balanced interim assessments, are medium-cycle and address intermediate goals on the way to meeting standards. The Smarter Balanced assessments are aligned to the standards, and any other interim or benchmark assessment used by districts or schools also need to be aligned to the standards. Typically administered quarterly or every six weeks, interim assessments cover a shorter period of instruction than long-cycle assessments and consequently give more detail about student learning. Results from interim assessments provide periodic snapshots of student learning throughout the year. These snapshots help teachers monitor how student learning is progressing and determine who is on track to meet the standards by the end

of the year and who is not, which may mean that a student is in need of additional support. When using or designing interim or benchmark assessments, teachers and school and district leaders determine reasonable expectations at various points in the year relative to meeting a CA CCSS for ELA/Literacy or CA ELD Standard at the end of the year. These interim goals for meeting the end-of-year standards at points along the school year are likely to look different than the end-of-year standard. Results from these assessments help teachers answer the following questions:

*Results from interim assessments provide periodic snapshots of student learning throughout the year. These snapshots help teachers monitor how student learning is progressing and determine who is on track to meet the standards by the end of the year and who is not, which may mean that a student is in need of additional support.*

- What have my students learned so far?
- Who has and who has not met intermediate goals?
- Who is and who is not on track to meet the standards by the end of the year?
- How are students performing on this test or assignment in those areas identified as weak on last year's California state long-cycle assessments?
- What are the strengths and areas of need in an individual's or groups' learning?
- Who are the students most in need of additional support? What do they need?
- What are the strengths and areas of need in my curriculum?
- What are the strengths and areas of need in my instruction?
- What improvements do I need to make in my teaching?

Administrators also use interim assessments to address many of these questions that are relevant to their decision-making needs, for example, programmatic, professional learning, and resource decisions.

If students are not making desired progress, teachers and administrators then consider whether changes are needed in curriculum and instruction while adjustments can still be made before the end of the year. Interim assessments also supply individual performance data. These data are useful in identifying individual student's strengths and learning needs. In addition, while these results sum up a period of learning, they can also be used formatively if steps are taken to respond to individual student's needs while there is still time within the year. In instances where no action is taken to support student learning, the results from these assessments remain summative only.

*If students are not making desired progress, teachers and administrators then consider whether changes are needed in curriculum and instruction while adjustments can still be made before the end of the year.*

Using data systems, including spreadsheets, interim assessment results are aggregated and displayed in graphs and charts, so teachers can identify patterns in their students' performance; interim assessment results are also disaggregated to provide information on the relative performance of individuals and subgroups. It is important that teachers and administrators have adequate professional learning and support to properly interpret the results of interim assessments so their conclusions and responses are appropriate and effective.

If districts, schools, or individual teachers use commercially-produced interim assessments, they need to consider technical quality to ensure that the assessments are appropriate for the intended purpose and that they are fully aligned with the CA CCSS for ELA/Literacy and CA ELD Standards. (See section on technical quality in this chapter.)

**Snapshot 8.5. Interim (Medium-Cycle) Assessment in Grade One**

All incoming first graders in a school are assessed at the beginning of the school year on the foundational skills of the CA CCSS for ELA/Literacy, specifically, print concepts, phonological awareness, phonics and word recognition, and fluency. Results from their end-of-year kindergarten assessment are used to determine which sections of the assessment are administered. For example, if a student's results indicate strong performance on a measure of print concepts, that part of the assessment is skipped, although close observations are made during class to confirm the previous year's assessments. The first-grade teachers find the results from the beginning-of-the-year assessment to be a useful starting point for their instructional planning, particularly because students may have either lost or made up ground during the summer. In addition, the teachers assess, or obtain help to assess, the primary language foundational literacy skills of their ELs who are new to the school and use this information for instructional decision-making.

After these initial assessments and implementation of appropriately designed instruction, students are administered interim foundational skills assessments every six weeks to determine progress. While the teachers regularly use formative assessment practices during their instruction to gather evidence of students' skill development and to adjust instruction accordingly, they use the results of the interim assessments to gauge the overall progress of individuals and the class as a whole, and to provide information regarding needed improvements in their teaching to ensure greater progress. The teachers also use the results as a means to evaluate and corroborate their own judgments about students' skill development in the period between the interim assessments' administration.

## *Assessing ELD Using Medium-Cycle Evidence*

As with all medium-cycle assessment, assessing progress in English language development using interim/benchmark assessments or classroom summative (such as end-of-unit/quarter) assessments should not take priority over short-cycle formative assessment. Short-cycle ELD formative assessment is assessment *for* learning and allows immediate teaching moves that support student learning *as it occurrs*. Medium-cycle ELD assessment is assessment *of* learning that has already occurred. It is typically not useful for providing immediate instructional support to students because the assessment evidence is too removed in time from the learning and is often too general.

Medium-cycle ELD assessments, such as classroom summative (including interim/benchmark) assessments, are useful, however, for evaluating a student's progress. They help teachers reflect on their instructional planning and implementation and make within-year program design and instructional adjustments. They also help school and district leadership identify or adjust professional learning and instructional materials decisions.

Periodic progress monitoring helps teachers determine the status of EL students' achievement of unit and intermediate (e.g., within-unit, quarterly) goals toward meeting particular CA ELD Standards as they progress through English language proficiency levels. In addition, monitoring helps teachers determine if students are advancing in English language proficiency or if they are *stalled* in particular areas. In other words, monitoring helps teachers know if their students are

*Periodic progress monitoring helps teachers determine the status of EL students' achievement of unit and intermediate (e.g., within-unit, quarterly) goals toward meeting particular CA ELD Standards as they progress through English language proficiency levels.*

*on track* for achieving end-of-year learning goals, differentiated for ELs using the CA ELD Standards, so that within-year instructional adjustments and refinements can be made. For example, a fifth-grade teacher examines a quarterly narrative writing task the whole class completes and uses the CA ELD Standards to analyze how EL students expand and enrich ideas in noun phrases (ELD.PI.5.4). Using this approach, she monitors how an individual EL student progresses throughout the year. A student who began the school year at an early Expanding level of proficiency, for example, might progress through the Expanding level (across the narrative writing samples) and, potentially, into the early Bridging level by the end of the year.

*Interim/benchmark assessments should be used judiciously and intentionally. . . . Teachers . . . ensure that tests match teaching and learning goals and that valuable instructional time is not sacrificed to administer and analyze tests—or any other type of medium-cycle assessment—that are not critical for monitoring ELD progress.*

Similarly, high school teachers design a two-month unit of study with a culminating, curriculum-embedded argument writing task. This writing task provides useful medium-cycle ELD assessment evidence if the writing is analyzed for degree of attainment of the learning goals tied to particular CA ELD Standards (e.g., how students are using verb types and tenses, organizing their writing, expanding noun phrases). This analysis helps teachers identify patterns in student learning outcomes (e.g., many students may need support in linking ideas throughout a text to create cohesion) as well as individual needs. Using the results of this analysis teachers plan further instruction, this within-year adjustment supports students' progress toward end-of-year goals in ELD.

Interim/benchmark assessments should be used judiciously and intentionally. Authentic classroom learning tasks, rather than multiple-choice tests or decontextualized performance tasks that focus on discrete grammatical skills and vocabulary knowledge, best inform ongoing teaching and learning in ELD. Teachers of ELs approach all assessments with a critical eye to ensure that tests match teaching and learning goals and that valuable instructional time is not sacrificed to administer and analyze tests—or any other type of medium-cycle assessment—that are not critical for monitoring ELD progress.

## Long-Cycle Assessment

Yearly assessments, such as the Smarter Balanced Summative Assessments, are long-cycle assessments. They typically assess students' mastery of standards at the end of the grade and provide student achievement results at several levels, including individual, school, district, and state. They sum up achievement after a year of learning and are therefore most appropriately used by schools and districts to monitor their own annual and longitudinal progress and to determine that individual students and groups of students are on track academically. The CELDT serves similar purposes with respect to measuring ELs' progress toward learning English. Schools and districts ensure that students in dual language programs are making steady progress toward biliteracy by including assessments in the relevant non-English language.

*They [long-cycle assessments] sum up achievement after a year of learning and are therefore most appropriately used by schools and districts to monitor their own annual and longitudinal progress and to determine that individual students and groups of students are on track academically.*

Long-cycle assessments also help teachers answer the following questions:

- What did my outgoing class of students learn? Did they meet the standards I was teaching them?
- What did my incoming class of students learn from last year to this year? Which standards did they achieve, and which did they not achieve?
- What are the overall strengths and areas of need in my class's learning?
- What are the strengths and areas of need in individual's and groups' learning?
- What are the strengths and areas of need in my curriculum?
- What are the strengths and areas of need in my instruction?
- Have the improvement strategies I/we put in place worked?

With data systems, assessment results are aggregated so that individual teachers and schools can look for patterns in their students' performance. They are also disaggregated to provide information on the relative performance of subgroups and the performance of individual students. School and district administrators use these assessment results to determine which students have and have not met the standards and identify the relative strengths and areas of need in curricula and programs. Long-cycle assessment results should be examined and discussed by teams of educators who sensitively analyze outcomes, responsively adjust instructional programs, plan professional learning, collaborate, and teach.

*Long-cycle assessment results should be examined and discussed by teams of educators who sensitively analyze outcomes, responsively adjust instructional programs, plan professional learning, collaborate, and teach.*

Long-cycle assessment results are appropriately used for system monitoring and accountability; reporting to parents on their individual child's achievement; adjustments to programs, curricula and instruction for the following school year; teachers' reflection on their instructional practices; and identifying teachers' professional learning needs. As indicated, results also provide a starting point for students' teachers in the following school year, offering a picture of classroom, subgroup, and individual strengths and weaknesses. Snapshot 8.6 provides a glimpse of these uses of long-cycle assessment.

### Snapshot 8.6. Long-Cycle Assessment in Grade Eight

Just before the new school year starts, eighth-grade English teacher Ms. Flora and her eighth-grade colleagues examine their incoming students' seventh-grade summative ELA assessment results to anticipate their students' learning needs. At the same time, they examine the prior year's CELDT results for their incoming EL students, some of whom have been in U.S. schools for only a couple of years and others for many years, as well as available data about their literacy proficiency in their primary language. The teachers want to make sure that they use all available information to design appropriately differentiated instruction for their students.

Last year's results suggest students may need considerable support in several areas, including close and analytic reading skills with respect to literature and informational text and writing effective arguments. To address weaknesses evident in the seventh-grade summative assessment results, Ms. Flora pays particular attention to the grade eight literature standards: (1) Cite textual evidence that most strongly supports an analysis of what the text says explicitly as well as inferences drawn for the text (RI.7.1), and (2) Compare and contrast the

**Snapshot 8.6. Long-Cycle Assessment in Grade Eight (cont.)**

structure of two or more texts and analyze how the differing structures of each text contribute to its meaning and style (RL.8.5). She focuses on the parallel standards for informational text as well. In addition, to address the weaknesses evident in the seventh-grade writing results, she works with her students extensively on the following standard: Write arguments to support claims with clear reasons and relevant evidence (W.6–8.1).

When she examines her students' eighth-grade ELA summative assessment results at the end of the year, the first question she considers is whether her students met the standards she identified as in need of considerable instructional attention. She is pleased to note that *most* students achieved proficiency on the targeted reading and writing standards. She is satisfied with the overall results and feels that the instructional focus that she and her colleagues identified for the year yielded positive results. However, some students did not meet the proficient levels on the state assessment, so she plans to follow up with her colleagues to examine the data to determine if there are students in other classes who did not achieve the standards. She also plans to closely view the data to see where specific areas of need lie and whether the results of summative assessment are consistent with what she observed through formative assessment and interim assessments. For her EL students, she plans to view EL students' results in light of their eighth-grade summative CELDT assessment results and note any relevant findings. This information provides evidence to help guide any changes in her instruction for next year's eighth graders. She also knows that her careful analyses will be valuable information to pass on to the ninth-grade teachers.

## *Ensuring Accessibility for ELs on Long-Cycle Assessments*

To ensure an accurate view of ELs' learning status, designated assessment supports may be needed. The intent is not to give EL students an unfair advantage over those who do not receive such support (Abedi and Ewers 2013). Rather, the goal of a support is to make an assessment more accessible for ELs and to allow students to demonstrate what they know and can do, thereby leveling the playing field and strengthening the validity of assessment results for ELs. The following factors should be considered when selecting assessment supports for ELs:

*Rather, the goal of a support is to make an assessment more accessible for ELs and to allow students to demonstrate what they know and can do, thereby leveling the playing field and strengthening the validity of assessment results for ELs.*

- Effectiveness: A support is effective in making an assessment more accessible to the recipients.
- Validity: A support does not alter the focal construct being assessed, that is, the outcomes of supported and unsupported assessments are comparable.
- Differential Impact: A support is sensitive to students' background characteristics and their academic standing; one size may not fit all.
- Relevance: A support is appropriate for the recipients.
- Feasibility: A support is logistically feasible to implement in the assessment setting (Abedi and Ewers 2013, 4).

The Smarter Balanced Assessment Consortium offers universal embedded online tools that improve the accessibility for all students, several embedded designated supports that improve accessibility for ELs, and accommodations for students with disabilities as required by their individualized education program (IEP) or 504 plan (Smarter Balanced 2013b). Examples of designated supports, depending on the type of assessment, include bilingual glossaries, translated test directions, and text-to-speech features. The type of support useful to ELs varies depending on the student's age and level of English language proficiency, as well as the subject area assessed, type of assessment task, and other factors. (See Usability, Accessibility, and Accommodations Guidelines from Smarter Balanced 2014 located at http://www.smarterbalanced.org/wordpress/wp-content/uploads/2014/08/SmarterBalanced_Guidelines.pdf.)

## Additional Methods of Medium- and Long-Cycle Assessment

Additional methods for evaluating student achievement in medium or long cycles include rubrics and student portfolios.

### *Rubrics*

Performance assessments that require students to demonstrate learning through an oral, written, or multimodal performance task (e.g., a presentation, a report) can be evaluated according to a rubric. A commonly accepted definition of a rubric is a document that articulates the expectations for an assignment by listing the criteria, or what counts, and describing levels of quality (Andrade, and others 2009). Criteria relate to the learning that students are being asked to demonstrate rather than the tasks themselves, and they should provide clear descriptions of performance across a continuum of quality (Brookhart 2013). The criteria are linked to standards and reflect what is required to meet a specific standard or cluster of standards.

Descriptions of performance are usually presented within score levels, and the number of score levels depends on the extent to which criteria across the levels can distinguish among varying degrees of understanding and skills. The knowledge and skills at one level differ distinctively from those at other levels (Lane 2013). Commercially produced performance assessments used for high stakes assessment purposes (e.g., placement or end-of-year grades) should provide evidence of their technical quality. (See the section on technical quality in this chapter.) Examples of evidence include review by language and literacy experts, review to ensure cultural and language sensitivity, and field tests that demonstrate that the rubric differentiates performance across levels of the rubric and across grades.

For classroom assessment, in situations where stakes are not so high, teachers—sometimes in collaboration with students—can develop rubrics for performance assessments. Co-construction of rubrics with students is a powerful way to build student understanding and acceptance of expectations. When creating rubrics, three points are important. First, rubrics should express as clearly and concisely as possible the expected performance at each level. Therefore, it is important to avoid ambiguous language. Before using the rubric, the language of the rubric is explained to students. Second, expectations are communicated through non-pejorative descriptions of what performance looks like at each level, reflecting a growth mindset. Third, the gradations of quality are specifically articulated across levels. Figure 8.6 presents an example of a rubric for scoring an essay. The dimensions of the rubric are listed on the left-hand side and the criteria are clearly described across four levels of performance along the top.

*Performance assessments that require students to demonstrate learning through an oral, written, or multimodal performance task (e.g., a presentation, a report) can be evaluated according to a rubric.*

***Figure 8.6. Essay Scoring Rubric***

| Dimensions | 4 | 3 | 2 | 1 |
|---|---|---|---|---|
| **Ideas and Content** | The essay has a clear thesis and supports it with evidence. Relevant comparisons between the paintings are made. Reasons for the similarities and differences are discussed in terms of the influence of one art movement on another. | The essay has a clear thesis. Comparisons between the art works are made. The discussion of influences might be thin. | An opinion is given. The support for it tends to be weak or inaccurate. May get off topic. | The thesis and support for it is buried, confused and/or unclear. |
| **Organization** | The paper has an interesting beginning, developed middle, and satisfying conclusion in an order that makes sense. Paragraphs are indented, have topic and closing sentences, and main ideas. | The paper has a beginning, middle and end in an order that makes sense. Paragraphs are indented; some have topic and closing sentences. | The paper has an attempt at a beginning and/or ending. Some ideas may seem out of order. Some problems with paragraphs. | There is no real beginning or ending. The ideas seem loosely strung together. Poor paragraph formatting. |
| **Voice and Tone** | The writing has a clear perspective, sophisticated style, and appropriate tone. | The style and tone are appropriate. The writer's perspective fades in and out. | The writer's perspective is obscure. The paper shows little awareness of audience and purpose. | The writing is flat, lacks a perspective, and uses an inappropriately formal or informal style and tone. |
| **Word choice** | The words used are descriptive but natural, varied and vivid. | The words used are correct, with a few attempts at vivid language. | The words used are ordinary. Some may sound forced or clichéd. | The same words are used repeatedly, some incorrectly. |
| **Sentence Fluency** | Sentences are clear, complete, begin in different ways, and vary in length. | Mostly well-constructed sentences. Some variety in beginnings and length. | Many poorly constructed sentences. Little variety in beginnings or length. | Incomplete, run-on and awkward sentences make the paper hard to read. |
| **Conventions** | Spelling, punctuation, capitalization, and grammar are correct. Only minor edits are needed. | Spelling, punctuation and capitalization are usually correct. Some problems with grammar. | There are enough errors to make the writing hard to read. | The writing is difficult to understand because of errors. |

**Source**
Andrade, Heidi G. 2013. *Essay Scoring Rubric.* Unpublished document.

It is preferable for teachers to design rubrics collegially as a group rather than as individuals. Taking advantage of how school teams already work together and ensuring that appropriate content expertise is represented are useful operating procedures for rubric development (Brookhart 2013). There is no specified frequency with which teachers should use rubrics. The use of a rubric depends on its purpose (Brookhart 2013). For example, a rubric may be used at regular intervals during a writing assignment or once each week to assess oral reading. Given the time and effort required to develop quality rubrics, it is important to identify learning goals or standards that are best assessed by performance tasks and rubrics, so that the investment in their development is worthwhile (Arter and Chappuis 2006).

*Rubrics can improve student performance, as well as monitor it, by making teachers' expectations clear and by showing students how to meet these expectations.*

Rubrics can improve student performance, as well as monitor it, by making teachers' expectations clear and by showing students how to meet these expectations. When teachers provide an evaluation of student work using a rubric, it should be clear to students what they need to do to improve in the future. Rubrics also support student self- and peer assessment. (See the section in this chapter on student involvement for more information on self- and peer assessment).

Rubrics are particularly useful for assessing oral language development, particularly for ELs. For example, rubrics can focus teachers' attention on particular discourse practices, grammatical structures, and vocabulary as they observe and listen to students' collaborative discussions, oral presentations, and informal conversations. These observations then guide instructional decision-making, including ways to structure conversations and productive group work, how to model different uses of English, and how to provide ample exposure to rich oral language, including from peers. The CA CCSS for ELA/Literacy and the CA ELD Standards can be used to design rubrics to gauge students' progress in oral language (including vocabulary and presentations), collaborative discussions, writing, and other areas of the curriculum. Given the interrelated nature of the two sets of standards, teachers create integrated rubrics that use both sets of standards, as well as the standards' companion appendices and documents, to avoid creating multiple rubrics for the same tasks.

## Portfolios

Student portfolios are another useful source of evidence for making judgments about student learning. A portfolio is a systematic collection of student work and related materials that tells the story of a student's activities, progress, and achievement in a given subject area (Arter and Spandel 1992; Venn 2000). Portfolios can provide a progressive record of student growth, or they can be used to demonstrate mastery of specific learning goals and contain samples only of a student's highest achievement (Venn 2000). Portfolios are considered either medium- or long-cycle assessments, depending on the length of the learning period covered.

Whatever the purpose of the portfolio, sufficient samples related to specific learning goals should be included to enable an evaluation of either growth or achievement (Chappuis, and others 2012). The specific learning goals are aligned to standards, and the evidence included in the portfolio reflects either students' progress toward meeting standards or achievement of specific standards.

Portfolios can include a range of evidence: student learning goals; samples of written work; images of work samples (e.g., digital images of models or other representations); audio samples (e.g., student narratives, oral presentations, or read alouds), video files (e.g., student performances, ASL-signed presentations); student reflections; teacher observations; teacher-student conference notes; and documentation of any other assessment results. Digital portfolios allow students to assemble and publicly present their work. Assembling a portfolio directly involves students in selecting its contents

as well as reflecting on the reasons selections were made, what they represent, and what they show about students' learning (Arter and Spandel 1992; Chappuis, and others. 2012).

The following are questions teachers should keep in mind when using portfolios:

- How representative is the work included in the portfolio of what students can really do?
- Do the portfolio pieces represent coached work, independent work, or group work?
- Do the portfolio pieces represent student language and literacy progress across the content areas?
- How well do the portfolio items match standards?
- Are there clear criteria for judging the work and do the criteria represent the most relevant dimensions of student work products?
- Is there a method for ensuring that evaluation criteria are applied consistently and accurately? (Arter and Spandel 1992)

*Assembling a portfolio directly involves students in selecting its contents as well as reflecting on the reasons selections were made, what they represent, and what they show about students' learning*

Well-developed criteria are used to evaluate portfolio items and establish the scoring process (e.g., the number of raters, when scoring takes place). It is also important to communicate whether the portfolio is to be rated as a whole or as individual samples, and if so, how the items are weighted. For example, are video performances of students' spoken language weighted more or less than their written artifacts?

Portfolios also provide valuable information about student progress to parents, particularly the parents of ELs and other language-minority students who may not be completely familiar with U.S. schooling practices and systems. Portfolios designed to *tell the story* of student growth during a particular time frame communicate to parents how their children are developing in a variety of areas valued by the standards and curricula. This information can help parents support their students' continued development and expand opportunities for collaboration between schools and families.

# Student Involvement

Whatever the assessment cycle, one goal of assessment is to promote a positive orientation to learning for students. Assessment, particularly when stakes are attached, creates a strong reason for learning. However, assessment can also impact the learner's willingness, desire, and capacity to learn (Harlen and Deakin Crick 2002). For example, if passing the test becomes the reason for learning, then students run the risk of developing a performance orientation, rather than a learning and mastery orientation (Ames and Archer 1988; Dweck 1999, 2006). Students with a performance orientation tend to use passive rather than active learning strategies, they avoid learning challenges, and their learning tends to be shallow rather than deep (Crooks 1988; Harlen and James 1997).

*Whatever the assessment cycle, one goal of assessment is to promote a positive orientation to learning for students. Assessment, particularly when stakes are attached, creates a strong reason for learning.*

While teachers can help students learn, only students can actually do the learning. For this reason, successful achievement of standards requires students to develop a learning orientation evidenced by an interest in learning and meeting challenges, and a belief that effort, engagement in learning, and the development of learning strategies lead to increased achievement.

If students are involved in the assessment process, they are more likely to develop a learning orientation than if they are solely passive recipients of test scores. They are also more likely to develop skills in setting goals, managing

the pursuit of those goals, and self-monitoring—all important 21st century skills (NRC 2012). Active student involvement in the assessment process is vital in developing student self-direction in learning. Crucial to student involvement in assessment, feedback is a critical factor in developing students' insight into their own learning and understanding (NRC 1999; OECD 2005).

## Feedback

Feedback indicates to students what they have done well—the degree to which they have met learning goals—and what they can do next to improve their learning (Bangert-Drowns, and others 1991). Importantly, feedback from teachers or peers should focus on tasks, processes students use, and students' self-regulation, rather than on students themselves (Kluger and DeNisi 1996; Hattie and Timperley 2007). Feedback, especially peer feedback, should avoid making comparisons with other students (Black and Wiliam 1998; Wiliam 2007). As Wiliam (2011) suggests, feedback should prompt a *cognitive* reaction (in which the learner focuses on active steps to achieve mastery) and not an *emotional* reaction (in which the learner experiences anxiety or embarrassment).

*Feedback indicates to students what they have done well—the degree to which they have met learning goals—and what they can do next to improve their learning.*

Long- and medium-cycle assessments usually produce a score indicating the status of achievement. While the scores typically tell students *what* they have achieved, they do not tell them *how* or *why* they achieved what they did. The role of teacher feedback in relation to these types of assessment results is to help students understand where they were successful or not and to set goals with students that inform them about where and how they need to improve. This approach requires that teachers spend time with students discussing assessment results and setting goals and strategies for improvement. Even when teachers use rubrics and provide evaluative scores, students still need feedback about how to improve. Although time consuming, the benefit for students is more assessment transparency and increased goal orientation and ownership of future learning.

When considering feedback to give EL students on their developing English language use, teachers should focus first and foremost on effective communication and meaning making. They take note of language resources (e.g., vocabulary, grammatical structures, discourse moves) students employ and plan ways with students to increase their use. Teachers encourage EL students to take risks when using English and establish a safe and supportive environment in which students are free to make *mistakes*—that are in fact normal developmental steps—in approximating complex academic uses of advanced English. For example, a student might ask, "How fast the lava go?" If a teacher stops to correct the student's grammar (e.g., to tell the student they must use the word *does*), the focus on meaning can be lost and the student may be discouraged from taking further risks. Instead, teachers think carefully about when and how to provide feedback on particular aspects of students' language use, including grammatical structures, vocabulary, and register. The teacher may at that moment simply acknowledge the student's question and recast the statement, thereby providing implicit feedback ("That's a great question! How fast does the lava flow? Let's read to find out."). In this example, the teacher also writes the recast question on a chart or using a document camera to provide a visual reinforcement of the oral modeling, and referring to a list of questions students have generated, the class chorally reads them together. In addition, the teacher notes what the individual student said and makes plans to address the grammatical structure of questions more explicitly during designated ELD. These examples do

*When considering feedback to give EL students on their developing English language use, teachers should focus first and foremost on effective communication and meaning making.*

not imply that explicit attention to English language development should not occur during content instruction. Rather, the way in which feedback is provided is carefully considered to maximize student meaning making and risk taking. Overcorrection, particularly when it feels to the student like ridicule (e.g., constantly requiring a student to repeat utterances in grammatically correct and complete sentences or chastising a student for not using standard English pronunciation), is not effective feedback. Overcorrection detracts from content knowledge development and discourages EL (and non-EL) students from participating in conversations and writing their ideas freely, thus impeding their language development.

> *. . . the way in which feedback is provided is carefully considered to maximize student meaning making and risk taking.*

**Snapshot 8.7. Student Involvement in Assessment in Grade Four**

Miss Nieto, a fourth-grade teacher, has a discussion with each of her students about their reading scores from an interim assessment. In her meeting with Henry, she notes that he has done well on the items related to using explicit details about the text and summarizing central ideas and is on track to meet the associated standards. She also discusses with Henry that his scores indicate that he is not as strong in using supporting evidence to justify or interpret how information is presented. Miss Nieto and Henry have a conversation about why he thinks he scored lower on these items. He tells her that he thinks he is beginning to better understand how to use evidence for justification but it continues to be difficult for him. She suggests that this should be something he consciously focuses on improving between now and the next interim assessment, and she provides some ideas to support his learning.

Feedback is particularly salient in the context of formative assessment. Students can receive feedback in three ways: from their teachers, from peers, and through their own self-assessment. The purpose of the feedback is to close the gap between the student's current learning status and the lesson goals (Sadler 1989). It is critically important that students be given opportunities to *use* the feedback, otherwise it does not serve the intended purpose.

## Teacher Feedback

Three categories of questions provide a frame for feedback to students in formative assessment (adapted from Hattie 2012, 130). The questions are crafted from the students' perspective with the aim of building students' involvement and ownership of learning through the formative assessment process.

1. Where am I going? What are my goals?
2. Where am I now? What progress am I making towards the goal?
3. Where to next? What do I need to do to make better progress?

To answer the first questions, both teachers and students need to be clear about the goal or target of the learning and what constitutes successful performance of learning. Answering the second set requires teachers and students to elicit and interpret evidence of learning. In other words, they need to decide where students' learning currently stands in relation to the learning goal. Answering the third set of questions guides students to take next action steps toward meeting the learning goal. Teacher feedback is required for students to answer the second and the third sets of questions. Teacher feedback indicates to students where they have been successful and provides hints or cues about what to do next.

**Snapshot 8.8. Teacher Feedback in Grade One**

Kathleen, a first grader, is preparing to read aloud to her teacher. Before she begins, Mr. Silverstein reminds her to think about the reading strategies they have been using. The text states: *Fish swim in the river*. Kathleen, reading very slowly, says: *Fish . . . swim . . . in . . . the . . . water. No. That's not water. It doesn't begin with 'w.' R* (says letter name) *r* (letter sound) *. . . i . . . v . . . River! Fish swim in the river.* Mr. Silverstein provides feedback after Kathleen finishes reading the sentence: *You did a very good job of using your decoding strategies to read the text accurately. Let's keep on reading and while you are reading think about whether what you are reading makes sense. It needs to! Also think about whether what you are seeing (that is, the letters and letter combinations) matches with what you are reading. You did that when you noticed that water could not be the right word. Water made sense, but the letters indicated a different, equally sensible word: river.*

## *Peer Feedback*

Peers are also sources of feedback for learning. Peer feedback has a number of advantages both for those students providing the feedback and those receiving it. It involves thinking about learning and can deepen students' understanding of their own learning. Research shows that the individuals providing feedback benefit just as much as the recipients because they are forced to internalize the learning goals and performance criteria in the context of someone else's work, which is less emotionally charged than their own (Wiliam 2006). The same three categories of questions listed in the teacher feedback section apply to peer feedback. Without clarity about the goal and the performance criteria, peers find it difficult to provide useful feedback to each other. Peers assess the status of classmates' learning against the same success criteria they use to check their own learning. Additionally, students need to learn to provide constructive feedback, so teachers should instruct and coach students on this as well. Notably, learners who become adept at giving and receiving feedback acquire valuable 21st century skills (NRC 2012).

*Peer feedback has a number of advantages both for those students providing the feedback and those receiving it. It involves thinking about learning and can deepen students' understanding of their own learning.*

**Snapshot 8.9. Peer Feedback in Grade Three**

In a third-grade class students are focusing on Speaking and Listening Standard 3.4, one of several that emphasize *presentation of knowledge and ideas*. Their learning goal is to write an informative speech to present to the class about a topic of interest to them. The criteria they need to bear in mind when writing their speeches include the following:

- Introduce your topic in a way that engages your audience.
- Put your ideas in a logical sequence.
- Make an impact on your audience with your ending.

Once students create an initial draft, they exchange their papers with a partner. Then students provide each other with feedback. One student's feedback to her partner is: *I liked how you started your speech with a question . . . that's a good way of getting your audience's attention. I think your ideas are logical. I think it would be a better impact at the end of your speech if you go back to your question and maybe finish with a sentence that tells how you answered the question.*

## Self-Assessment

Teacher and peer feedback are externally provided. When students are involved in self-assessment they are generating *internal feedback*. Generating and acting on internal feedback is a form of metacognition and self-regulation. Metacognition is thinking about one's thinking, and self-regulation refers to the ability of learners to coordinate cognitive resources, emotions, and actions in order to meet learning goals (Boekaerts 2006). In the realm of 21st century learning, metacognition and self-regulation are important skills (NRC 2012), and the most effective learners are self-regulating (Butler and Winne 1995; Pintrich 2000; Schunk and Zimmerman 2008). Additionally, teaching students metacognition raises their performance (e.g., Lodico, and others 1983) and helps them generalize and transfer what they learn to new situations (Hacker, Dunlosky, and Graesser 1998). Because of the importance of metacognition and self-regulation for successful learning, teachers ensure that students develop these skills in the context of language and literacy learning.

*When students are involved in self-assessment they are generating* internal feedback. *Generating and acting on internal feedback is a form of metacognition and self-regulation.*

Self-assessment can be developed from the early grades onwards (Perry, and others 2002; Puckett and Diffily 2004). For example, a first-grade teacher provides her students with a graphic organizer with the headings: *date, book title, my goal today as a reader, pages read, how well did I meet my goals?* She asks students to set goals for their daily independent reading time and, at the end of the session, to think about how well they met their goals. During her weekly individual reading conferences with students, she reviews the self-assessment sheets, and when a student has not met the goal the teacher asks what he or she did or needs to do to improve. Together, they identify a strategy for the student to use. In addition to providing the students with opportunities for self-assessment, the teacher offers advice on strategies for improvement, which in turn become part of students' internal repertoire of strategies that they can employ on subsequent occasions. In effect, they develop the skills of self-regulation.

Self-assessment becomes more sophisticated as students gain more experience. For example, in a ninth-grade science class in which the teacher integrates ELA and science standards, the students are involved in a short research project on distinct regions of the brain. As called for in the ELA writing

standards for literacy in science (WHST.9–10.6), they are to display their information "flexibly and dynamically." Students in this class have time toward the end of every session to complete a reflection and planning log and answer the following questions: *What was successful about your learning today? What difficulties or problems did you encounter? How did you manage those difficulties? Were you successful? If not, what plans do you have for dealing with them in the next lesson?* These logs serve as a means of self-assessment for students and support self-regulation since students have to generate strategies to solve difficulties. The logs are also sources of information for teachers about the progress students are making on their projects.

# Assessment of ELD Progress

Assessing ELD progress, particularly the development of *academic uses of English* in each discipline, is a responsibility shared by all educators in schools and districts where ELs are students. (See chapter 11 in this *ELA/ELD Framework* for information on district and school leadership responsibilities for monitoring ELD progress.) Districtwide and schoolwide assessment and professional learning *systems* are critical for ensuring EL students' achievement of the overarching goals of ELA/literacy and ELD instruction: students develop the readiness for college, careers, and civic life; attain the capacities of literate individuals; become broadly literate; and acquire the skills for living and learning in the 21st century. (See the outer ring of figure 8.1). However, *teachers* are the ones who ultimately ensure that every day, each of their EL students has full access to grade-level curricula and that they develop academic English in a timely—and even accelerated—manner. Teachers' deep understandings of the CA CCSS for ELA/Literacy, the CA ELD Standards, and other content standards are critical to effective assessment for and of learning because these standards guide instructional and assessment practices with ELs.

*Teachers' deep understandings of the CA CCSS for ELA/Literacy, the CA ELD Standards, and other content standards are critical to effective assessment for and of learning because these standards guide instructional and assessment practices with ELs.*

The CA ELD Standards provide outcome expectations at different English language proficiency levels (Emerging, Expanding, Bridging) so that teachers can differentiate their instruction according to individual EL students' language learning needs on particular standards. Because the CA ELD Standards delineate proficiency levels which EL students are expected to progress through during the year (and in fact, they may progress through more than one level in a single school year), teachers carefully attend to the ELD progress of their EL students on a frequent and ongoing basis. As described previously in this chapter, this ongoing monitoring of student progress involves using short-cycle formative assessment (minute-by-minute, daily, weekly), as well as medium-cycle assessment for formative purposes (monthly, end-of-unit, interim, benchmark, and other periodic time frames). Attending to the developing capacities and emerging or persistent needs of ELs is consistent with the assessment approaches teachers employ for all students. However, because ELs are learning English as an additional language *at the same time as* they are learning content knowledge through English (and therefore have particular English language learning needs), teachers take additional steps to assess ELD progress and act on evidence gathered from assessment. They consider the following questions:

- How do I determine what my EL students' levels of English language proficiency (Emerging, Expanding, Bridging) are on different CA ELD Standards?

- How can I use information about my students' English language proficiency levels on different CA ELD Standards, as well as other relevant information, to design and provide targeted instruction that fosters language-rich learning opportunities?
- How often should I assess ELD progress? Which kinds of evidence-gathering approaches and tools are most appropriate for different purposes?
- How will I know if my EL students are making sufficient progress in developing English on a daily or weekly basis and over longer periods of time?
- How can I include my EL students in assessing their own ELD progress and support them to be conscious of and intentional in their English language learning?

Guidance for addressing these questions follows. This guidance focuses on how classroom teachers—including ELA teachers, teachers in other content areas, ELD teachers, and EL specialists who support content teachers—can use the CA ELD Standards to assess the ELD progress of each of their EL students. The CA ELD Standards support teachers' formative assessment practices by offering descriptions of what EL students can be expected to do at the end of each English language proficiency level (Emerging, Expanding, Bridging). These expectations help teachers focus their formative assessment practices as EL students use English while learning content, gauge their EL students' developing capacities in English, and adjust instruction and learning opportunities. The CA ELD Standards also support teachers in their assessment of learning and to use this information for formative purposes. The examples offered here are intended to be used *in addition to—not instead of*—those provided in the rest of this chapter.

## Assessing ELD Progress in Writing

One way teachers can observe and respond to their EL students' development of written language is by using a *language analysis framework for writing*, based on the CA ELD Standards and aligned to teachers' learning goals and success criteria for writing. A language analysis framework allows teachers to observe and analyze student language in *linguistic terms* with more specificity than is often found in ELA rubrics or other tools for evaluating writing. For example, feedback to students on writing, such as "interesting beginning, developed middle, and satisfying conclusion," "could use more varied sentence patterns," or "needs some colorful vocabulary" may be sufficient for some students to improve their writing. However, this feedback may not be explicit enough for many students, including ELs, to act on (Fang and Wang 2011). Teachers providing this type of feedback may know intuitively what kind of writing they would like to see their students produce, but without specific feedback on the *language resources* that constitute "varied sentence patterns" or "colorful vocabulary," their feedback is elusive to ELs, and such language use remains a "hidden curriculum" (Christie 1999).

A language analysis framework for writing, drawing from the CA ELD Standards and other resources focused on language development, helps teachers provide a level of explicitness about the *specific language resources* that students can use in their academic writing to meet identified learning goals and success criteria in different disciplines. A language analysis framework provides *framing questions* that students can ask themselves as they are writing and as they examine writing. Guidance for composing and revising their writing can help students structure their texts cohesively and use expected grammatical structures and vocabulary. Explicitly focusing on language makes expectations for writing more transparent. Teachers can also use a language analysis framework to determine how well students use particular language resources in a piece of writing to provide useful feedback to students and adjust instruction accordingly. An example of a language analysis framework for writing in the upper elementary grades, developed using the CA ELD Standards and their English language proficiency descriptors (CDE 2014), as well as research on language development, is provided in figure 8.7.

***Figure 8.7. Language Analysis Framework for Writing***

| Language Analysis Framework for Writing | | | | |
|---|---|---|---|---|
| **Content Knowledge and Register** | **Text Organization and Structure** | **Grammatical Structures** | **Vocabulary** | **Spelling and Punctuation** |
| Is the overall meaning clear?<br>Are the big ideas there and are they accurate?<br>Is the text type (e.g., opinion, narrative, explanation) appropriate for conveying the content knowledge?<br>Does the register of the writing match the audience? | Is the purpose (e.g., entertaining, persuading, explaining) getting across?<br>Is the overall text organization appropriate for the text type?<br>Are text connectives used effectively to create cohesion?<br>Are pronouns and other language resources used for referring the reader backward or forward? | Are the verb types and tenses appropriate for the text type?<br>Are noun phrases expanded appropriately in order to enrich the meaning of ideas?<br>Are sentences expanded with adverbials (e.g., adverbs, prepositional phrases) in order to provide details (e.g., time, manner, place, cause)?<br>Are clauses combined and condensed appropriately to join ideas, show relationships between ideas, and create conciseness and precision? | Are general academic and domain-specific words used, and are they used accurately?<br>Are a variety of words used (e.g., a range of words for "small": little, tiny, miniscule, microscopic)? | Are words spelled correctly?<br>Is punctuation used appropriately? |

**Sources**

From

Spycher, Pamela, and Karin Linn-Nieves. 2014. Reconstructing, Deconstructing, and Constructing Complex Texts. In *The Common Core State Standards in English Language Arts/Literacy for English Language Learners: Grades K–5*, edited by Pamela Spycher. Alexandria, Virginia: TESOL Press.

As adapted from

Derewianka, Beverly. 2011. *A New Grammar Companion for Teachers*. Sydney, NSW: Primary English. Teaching Association.

Gibbons, Pauline. 2009. *English Learners, Academic Literacy, and Thinking: Learning in the Challenge Zone*. Portsmouth, NH: Heinemann.

Spycher, Pamela. 2007. "Academic Writing of English Learning Adolescents: Learning to Use 'Although.'" *Journal of Second Language Writing* 14 (4):238–254.

*Students also use a language analysis framework or related tools, such as a* success criteria *document addressing particular language areas, to evaluate and refine their own writing. Tools such as these support students to reflect on their work and ask themselves the same types of questions—either as they are writing or during a writing conference with peers or teachers—that teachers ask when analyzing student writing.*

Teachers use such a framework (adjusted appropriately for grade level or span), accompanied by knowledge of their students (including students' proficiency level on different CA ELD Standards), *for observing* what students are doing while writing and *for evaluating* students' written products. Having a framework for analyzing writing helps teachers focus on one or two areas to provide *just-in-time* scaffolding.

Students also use a language analysis framework or related tools, such as a *success criteria* document (Heritage 2014) addressing particular language areas, to evaluate and refine their own writing. Tools such as these support students to reflect on their work and ask themselves the same types of questions—either as they are writing or during a writing conference with peers or teachers—that teachers ask when analyzing student writing. Using a language analysis framework also helps students to monitor their own progress in writing.

The following annotated writing sample (figure 8.8) illustrates the use of a language analysis framework to analyze student writing to determine next steps for instruction. The example was written independently by an EL student in the fifth grade after several days of instruction during which students jointly constructed several short sections of a longer text on bats. (See vignette 5.3 in chapter 5 of this *ELA/ELD Framework* for an example of the task, text reconstruction). The students used their shorter reconstructed texts, along with other texts, to construct their own texts on bats at the end of the week. The teacher analyzed the writing for formative assessment purposes and to discuss refinements with students rather than for grading students' writing.

*Having a framework for analyzing writing helps teachers focus on one or two areas to provide* just-in-time *scaffolding.*

## *Figure 8.8. Student Annotated Writing Sample Using the CA ELD Standards*

### Susana's Text

Bats

Bats are important because they eat mosquitos, insects, mice, frogs and other small animals that could become pest in your house. Tey are also important because the spreed pollen and seeds and because of that more plants grow. If it werent for bats we wouldnt have all the food that we have now.

There are over 1,200 speicies of bats. The largest bat is the flying fox and you could find it in Austrilia. 3 species of bats are bumble bee bat, fruit bat, and even vampire bats. Bat can be able to damage many plants. They dont suck blood like in horror movies. They drink it like little kittens. Tey are the only mamles that could fly.

Bats are in danger because people are scared of them. They are scared of them that they burn there homes. There are less bats know than they use! to be. In Austrilia flying foxes are dieing of heat waves. Ith 15 years over 30,000 bats are dying because of us.

**Summary Notes and Next Steps:**

Discuss with Susana:

- Ordering of the three chunks, need for introduction that foregrounds the chunks, conclusion that sums them up
- Review whether information in each chunk fits there and if ideas in each chunk could be expanded more
- Show where clauses are combined to show relationships between them (e.g., using *because*), and ask her to see where she could do the same to combine other clauses

Discuss with the class (based on patterns in other students' writing):

- how register shifts when *you*, *we*, *us* are used
- how connecting and condensing ideas (clause combining or other ways) creates relationships between ideas and reduces repetition (maybe a mini-lesson with examples from student writing we revise together)
- how to use text connectives (maybe revise a piece of writing together and add in text connectives where needed to create cohesion)

### Annotations

**Content and register:**

- Big ideas and lots of informative details provided, mostly accurate information
- Some information needs more clarity (bats aren't in danger just because people are scared of them)
- *You, we, us* is used (less formal register)

**Text structure and organization:**

- Organized logically into three chunks (*why bats are important, species of bats, why bats are in danger*)
- Some information doesn't seem to fit in the chunks (*bats damaging plants*)
- Missing an introduction and conclusion, order may not be logical
- Pronoun reference: *because of that* used accurately to condense and link to previous sentence (cohesion)
- Could use more text connectives (cohesion)

**Grammatical Structures**

- Some appropriate clause combining to link ideas and show relationships
- Some clause combining needs work (*They are scared . . . that they burn . . .*) and more could be used
- Phrases could be expanded to include more details about where, when, etc.

**Vocabulary:**

- Domain-specific (*mammals*, *species*, *pollen*) and general academic (*spread*, *damage*) vocabulary used accurately

**Spelling and punctuation:**

- Mostly accurate, with some approximations (*mamles, dieing*)

**Source**

Adapted from

Spycher, Pamela. 2007. "Academic Writing of English Learning Adolescents: Learning to Use 'Although.'" *Journal of Second Language Writing* 14 (4): 238–254.

Student text from

Spycher, Pamela and Karin Linn-Nieves. 2014. "Reconstructing, Deconstructing, and Constructing Complex Texts." In *The Common Core State Standards in English Language Arts/Literacy for English Language Learners: Grades K–5*, edited by Pamela Spycher. Alexandria, Virginia: TESOL Press.

Using a language analysis framework for writing also guides discussions about writing. For example, in a writing conference during which Susana has an opportunity to discuss her writing with her teacher, Susana's teacher opens the conversation by asking Susana to identify areas where she feels her writing needs refinement. Susana's teacher continues the conversation by acknowledging specific areas of strength (e.g., "I see that you are providing lots of great content information about bats and that you're organizing the information in a way that helps the reader follow your ideas."). She asks probing questions to prompt Susana to notice areas for refinement and draws attention to text that needs refinement. She uses the CA CCSS for ELA/Literacy for grade five to frame her learning goals for the conference and the CA ELD Standards to help her to provide targeted support, based on Susana's English language proficiency level. Some of the questions she asks to prompt Susana's thinking and extend her use and understanding of English are the following:

*Using a language analysis framework for writing also guides discussions about writing.*

- How could you orient the reader to what your paper is about? How could you let them know in advance about the categories you've chosen to include?
- Does all of this information belong together in this section?
- How could you expand this idea to add more detail?
- How could you combine these ideas to show the relationship between them?
- Is there another word or phrase that would help you get your meaning across in a more precise way?

After examining student writing, teachers determine whether and in what ways students have progressed, and what next instructional steps are needed to support further language learning. For example, if a group of EL students at the Emerging level of proficiency are not yet using pronouns to refer to information that has already been presented in a text, their teachers model how to do this, provide opportunities to apply this new language resource to their own writing, and continue to draw their attention to pronoun reference until students have internalized this understanding. If EL students at the Expanding level are already using pronoun reference but not yet using more sophisticated cohesive language resources, such as the use of demonstratives (e.g., *this*, *that*) or nominalization (e.g., *the result* of *environmental degradation* . . .), their teachers show them examples of these language resources in the texts they read, have students analyze texts and provide multiple opportunities for students to apply this awareness of how English works to their own writing. Teachers monitor how well students *take up* these language resources in their writing over time and provide targeted feedback to the whole class, small groups, or individual students so that they continue to progress in their English language development.

*After examining student writing, teachers determine whether and in what ways students have progressed, and what next instructional steps are needed to support further language learning.*

## Assessing ELD Progress in Oral Language

Oral language use is a critical component of English language development, and observing how students are developing the language skills, abilities, and awareness needed for collaborative conversations and other oral language tasks, such as oral presentations, is essential. Teachers carefully plan collaborative learning opportunities and intentionally observe their EL students as they engage in these tasks, so they can provide just-in-time scaffolding to advance students' oral language to higher levels of proficiency. These formative assessment practices, which should remain the top priority during classroom instruction, are complemented by more formal evidence-gathering strategies and tools for observing and documenting progress in English oral language development.

*Oral language use is a critical component of English language development, and observing how students are developing the language skills, abilities, and awareness needed for collaborative conversations and other oral language tasks, such as oral presentations, is essential.*

For example, in grade seven, students are expected to engage in small group discussions about complex texts. One of their conversations revolves around an informational science text they are currently reading. As the students discuss their ideas about the text and extend their thinking about the content, the focus of teachers' observations is primarily meaning making. In other words, the teachers look for evidence that students understand the content of the text, make appropriate inferences based on the text and their background knowledge, use relevant examples, and extend their understandings of the text by asking their peers questions and answering questions posed to them.

Teachers also observe how their EL students use English to convey their ideas and engage in academic conversations in the context of authentic, meaningful interactions about complex texts and topics. Teachers also observe their non-EL students' academic language development during these meaningful interactions with texts, tasks, and others. However, the CA ELD Standards specifically help teachers determine, by English language proficiency level (Emerging, Expanding, Bridging), the types of language resources their EL students should be able to use in collaborative conversations. This forms the basis for evidence-gathering strategies and tools that help focus observations and determine next steps for supporting students' oral language development.

Such strategies and tools are used to focus attention on specific language uses that teachers and students determine as areas of growth. Observation tools help teachers notice how their students are progressing in their capacity to engage in collaborative conversations. As teachers develop deeper understandings of the CA ELD Standards, they increasingly notice how their EL students are using English in the context of specific CA ELD standards. They also become more skilled at identifying where on the ELD continuum their students are and their next steps in developing their academic uses of language.

Observation tools should be used strategically and purposefully. For example, teachers might use a formal observation tool monthly or quarterly. The tool is used more frequently with some students (e.g., newcomer ELs at the early Emerging level) and less frequently (for students at the late Bridging level) because the tool is intended to complement the ongoing observations teachers make every day. Oral language observation tools are intended to provide information to teachers that informs their instructional decisions, not for awarding grades to students. Figure 8.9 provides an example of an observation tool for monitoring grade-seven EL students at different places along the ELD continuum as they use English in collaborative conversations.

***Figure 8.9. Grade Seven Collaborative Conversations Observation Notes***

| **Collaborative Conversations Observation Notes** | | | |
|---|---|---|---|
| ***English Language Development Level Continuum*** → **Emerging** → **Expanding** → **Bridging** → | | | Students said . . . (note students' names and comments) |
| **CA ELD Standards in Focus:** | | | |
| **Exchanging Ideas Respectfully** *(ELD.PI.7.1)* | | | |
| Engage in conversational exchanges and express ideas on familiar topics by asking and answering *yes-no* and *wh-* questions and responding using simple phrases. | Contribute to class, group, and partner discussions by following turn-taking rules, asking relevant questions, affirming others, adding relevant information, and paraphrasing key ideas. | Contribute to class, group, and partner discussions by following turn-taking rules, asking relevant questions, affirming others, adding relevant information and evidence, paraphrasing key ideas, building on responses, and providing useful feedback. | |
| **Supporting Opinions and Persuading Others** *(ELD.PI.7.3)* | | | |
| Negotiate with or persuade others in conversations (e.g., to gain and hold the floor or ask for clarification) using learned phrases (e.g., *I think . . ., Would you please repeat that?*) and open responses. | Negotiate with or persuade others in conversations (e.g., to provide counter-arguments) using learned phrases (*I agree with X, but . . .*), and open responses. | Negotiate with or persuade others in conversations using appropriate register (e.g., to acknowledge new information) using a variety of learned phrases, indirect reported speech (e.g., *I heard you say X, and I haven't thought about that before*), and open responses. | |
| **Connecting Ideas** *(ELD.PII.7.6)* | | | |
| Combine clauses in a few basic ways to make connections between and join ideas (e.g., creating compound sentences using *and, but, so*; creating complex sentences using *because*). | Combine clauses in an increasing variety of ways (e.g., creating compound and complex sentences) to make connections between and join ideas, for example, to express a reason (e.g., *He stayed at home on Sunday in order to study for Monday's exam*) or to make a concession (e.g., *She studied all night even though she wasn't feeling well*). | Combine clauses in a wide variety of ways (e.g., creating compound, complex, and compound-complex sentences) to make connections between and join ideas, for example, to show the relationship between multiple events or ideas (e.g., *After eating lunch, the students worked in groups while their teacher walked around the room*) or to evaluate an argument (e.g., *The author claims X, although there is a lack of evidence to support this claim*). | |
| **Quick Observation Analysis** | | | |
| **Next steps** | | | |

The tool provided in figure 8.9 is used to complement the more informal minute-by-minute observations teachers make of their students during collaborative conversations. Care should be taken in implementing such tools. For example, attempting to observe too many standards at once or using the tool too often can be frustrating and counter-productive. Teachers need support and the flexibility to use such tools in ways that best inform their instructional practice.

*Teachers should develop and employ assessment approaches that support the learning goals they have for all students and strategically select additional approaches (when needed) that help them ensure that their EL students are advancing along the ELD continuum in a timely manner.*

The approaches and tools for assessing ELD progress provided in the preceding pages illustrate how teachers can attend to their EL students' progress in developing English as an additional language. They are not meant to be prescriptive. Teachers should develop and employ assessment approaches that support the learning goals they have for all students and strategically select additional approaches (when needed) that help them ensure that their EL students are advancing along the ELD continuum in a timely manner.

Progress in ELD is also monitored through appropriate use of large-scale summative assessments, such as the CELDT (and in time, the ELPAC). As delineated in figure 8.3 in this chapter, such summative assessments are not intended for planning daily instruction. Rather, the evidence from large-scale summative assessments related to ELD helps schools and districts evaluate and adjust the design of instructional programs provided to ELs and measure ELs' progress in learning English *from year-to-year*. Systematic monitoring should determine if EL students are progressing in their English language development within appropriate time frames and employ clearly defined protocols for action if they are not.

For example, a school leadership team conducts a systematic and careful analysis of year-to-year ELD progress, based on current and previous years of summative assessment results (in concert with other measures of student achievement), to identify EL students who demonstrate the following:

- Readiness to reclassify as English proficient
- Progress in English language development at an appropriate rate
- Stalled progress in English language development

The team carefully identifies possible reasons for students' progress or the lack thereof. Using the results of their analyses, the team determines specific and timely next steps for instructing individual students, as well as appropriate adjustments and additions to program design, professional learning, and the school or district's comprehensive assessment system. In addition, the team ensures that an accountability system is in place to measure the efficacy of these adjustments and additions. Additional guidance on reclassification is provided in chapter 11 of this *ELA/ELD Framework*.

# Assessment for Intervention

Screening, diagnostic, and progress-monitoring assessments are discussed in this section. Screening assessments identify students who may have difficulties, diagnostic assessments give specific information about the difficulties, and progress-monitoring assessments provide feedback on whether planned interventions to address difficulties are working. These assessments operate in short or medium cycles.

## Universal Screening (Medium Cycle)

Universal screening is a critical first step in identifying students who may be at risk of experiencing difficulty with reading and who may need more instruction. Universal screening consists of brief assessments focused on target skills that are highly predictive of future outcomes (Jenkins 2003).

An expert panel convened by the U.S. Department of Education's Institute of Education Sciences recommended that screening take place at the beginning of each school year in kindergarten through grade two, with a second screening conducted mid-year in kindergarten and grade one (Institute of Education Sciences [IES] 2009).

*Universal screening is a critical first step in identifying students who may be at risk of experiencing difficulty with reading and who may need more instruction.*

The panel report recommends that the following target areas should be screened at each grade: Kindergarten screening batteries include measures assessing letter knowledge, phonemic awareness, and expressive and receptive vocabulary. As children move into grade one, screening batteries include measures assessing phonemic awareness, decoding, word identification, and text reading. By the second semester of grade one decoding, word identification, and text reading measures include speed as an outcome. Grade two batteries include measures involving word reading and passage reading. For a reasonably accurate identification of students, the report also recommends the use of two screening measures at each juncture. When schools or districts select screening measures they should carefully examine the technical information available from the publisher's manual (IES 2009).

## Diagnostic Assessment (Medium Cycle)

While the purpose of diagnostic assessments is to improve student learning, they should not be confused with short-cycle formative assessment. Formative assessment is used to guide ongoing decisions about student learning, whereas diagnostic assessment is used to identify areas where intervention may be needed to improve student learning (Carnegie Council on Advancing Adolescent Literacy 2010).

Poor performance might reflect any one of a number of problems including, but not limited to, struggles with language and literacy. For example, if students are having difficulty understanding grade-level text, they may have short-term memory issues, may not read fluently enough to focus their attention on meaning making, or may not be making connections across phrases and sentences in the text. Diagnostic assessment is the means by which to identify the precise source(s) of the student's difficulty so that an appropriate intervention can be planned. Timely identification of students' difficulties is essential to ensuring the right intervention is made so students can progress.

*Formative assessment is used to guide ongoing decisions about student learning, whereas diagnostic assessment is used to identify areas where intervention may be needed to improve student learning.*

Great care should be taken when approaching diagnostic assessment in English for ELs and students who are deaf. For example, an EL at the Emerging level of English language proficiency or a student who is deaf may appear to struggle with reading comprehension when reading a complex text in English. However, it could be that the student has not had sufficient opportunity to develop the language resources in English (including vocabulary and grammatical structures) or background knowledge needed to apply reading comprehension strategies. With appropriately adjusted instructional support, the students may demonstrate comprehension. Diagnostic assessments administered in English to ELs and students who are deaf need to be interpreted carefully. Teachers should consider possible linguistic and cultural biases of assessments

(see section on technical quality in this chapter), use multiple types of assessments (including, as appropriate, assessments given in the primary language) to gain a comprehensive portrait of a student's learning needs, and compare the student's results to those of her or his peers who are ELs or who are deaf and not just to native English speakers.

> *. . . any test that uses language is in part a test of language. Therefore, for ELs and students who are deaf and use American Sign Language, every test written in English—regardless of the content area—is partially a test of their English language proficiency and may not adequately assess their content area knowledge and skills*

According to the *Standards for Educational and Psychological Testing* (American Educational Research Association [AERA], American Psychological Association [APA], and National Council on Measurement in Education [NCME] 1999), any test that uses language is in part a test of language. Therefore, for ELs and students who are deaf and use American Sign Language, every test written in English—regardless of the content area—is partially a test of their English language proficiency and may not adequately assess their content area knowledge and skills (Abedi 2002). For this reason, it may be beneficial to assess them in their primary language in order to gain a more complete picture of their strengths and needs. However, it may not be appropriate to use content assessments in the primary language with every EL student. For example, students who are literate or receiving formal instruction in their native language in a content area and who are at lower English language proficiency levels are more likely to benefit from a content assessment in the primary language than those who are not (Pennock-Roman and Rivera, 2011; Bowles and Stansfield 2008; Stansfield and Bowles 2006). Similarly, evaluating emerging bilinguals' writing by looking at their Spanish writing side by side with their English writing can help teachers see the how the languages reinforce one another, and provide a more comprehensive view of the students' developing biliteracy (Soltero-Gonzalez, Escamilla, and Hopewell 2012).

A range of assessments is available for diagnosing the source of a student's difficulties, and it is important to ensure the appropriateness of these assessments for diagnostic purposes. (See the section on technical quality in this chapter). Because administering and interpreting some diagnostic assessments requires special training and licensure, when selecting diagnostic assessments it is important to determine if the school has access to professionals who are qualified to administer them. Teachers benefit from working closely with reading specialists who have the necessary specialized knowledge to interpret diagnostic data and provide guidance regarding specific interventions (International Reading Association 2000). It is advantageous for all available professionals (e.g., teacher, reading specialist, and school psychologist) to work together in diagnosing a student's difficulties and planning appropriate interventions (Joseph 2002).

> *It is advantageous for all available professionals (e.g., teacher, reading specialist, and school psychologist) to work together in diagnosing a student's difficulties and planning appropriate interventions.*

## Progress Monitoring (Short or Medium Cycle)

Progress monitoring (sometimes referred to as curriculum-based measurement or curriculum-based assessment) is the practice of assessing students' academic performance on a regular basis for three purposes: (1) to determine whether students are profiting appropriately from the instructional program, including the curriculum; (2) to create more effective programs for those students who are not benefitting; and (3) to estimate rates of student improvement (National Research Center on Learning Disabilities 2006). To implement progress monitoring, a student's current level of

performance is determined and goals are established for learning for a specific period. The student's academic performance is assessed on a regular basis (see IES 2009 recommendations discussed previously) and progress toward meeting the goal is determined by comparing the actual and expected rates of learning.

*To implement progress monitoring, a student's current level of performance is determined and goals are established for learning for a specific period.*

In addition to the general screening measures described previously, a system of progress monitoring is recommended in response to intervention (RtI) programs (IES 2009). Although these recommendations are grounded in research related to RtI, they are consistent with the more comprehensive Multi-Tiered System of Supports (MTSS) structure recommended by this *ELA/ELD Framework*. (See chapters 2 and 9.) Based on available evidence, the panel report recommends that progress-monitoring assessments be administered to Tier 2 students at least once each month. For those students who are not making sufficient progress, a Tier 3 intensive intervention should be planned. Progress-monitoring assessments are used in Tier 3 to determine the effectiveness of the intervention (IES 2009).

The National Association of State Directors of Special Education (NASDSE) identified nine essential characteristics of progress monitoring. Recommendations advise that progress monitoring assess marker variables that have been demonstrated to lead to the ultimate instructional target, be sensitive to small increments of growth over time, be administered repeatedly using multiple forms, be administered efficiently over short periods, and result in data that can be summarized in teacher-friendly data displays (NASDSE 2005, 25–26).

If teachers, schools, or districts wish to adopt progress-monitoring assessments, careful attention needs to be paid to the technical quality of any proposed assessments to ensure they are appropriate for the intended purpose. (See the section on technical quality in this chapter.)

# Mandated California Assessments

On October 2, 2013, Assembly Bill 484 established the California Assessment of Student Performance and Progress (CAASPP) assessment system, which replaces the Standardized Testing and Reporting (STAR) program. The primary purpose of the CAASPP system is to assist teachers, administrators, and students and their parents by promoting high-quality teaching and learning through the use of a variety of assessment approaches and item types.

*To ensure the assessments address the full range and depth of the CA CCSS for ELA/Literacy, and the breadth of achievement levels, Smarter Balanced Summative Assessments combine item types, including selected response (multiple-choice items with one or multiple correct responses and two-part items) and constructed response (students write a short text or long essay in response to a prompt).*

Beginning in the 2014–2015 school year, student performance in grades three through eight and in grade eleven is assessed by annual summative assessments developed by the Smarter Balanced Assessment Consortium and administered in accordance with CAASPP regulations, CCR Section 855(b)(1) and (2). See figure 8.10. This state law exempts ELs from taking the ELA portion of the SBAC assessment if they have been enrolled in a U.S. school for less than 12 months.

To ensure the assessments address the full range and depth of the CA CCSS for ELA/Literacy, and the breadth of achievement levels, Smarter Balanced Summative Assessments combine item types, including selected response (multiple-choice items with one or multiple correct responses and two-part items)

and constructed response (students write a short text or long essay in response to a prompt). For example, for the third-grade reading standard, *determine the main idea of a text; recount the key details and explain how they support the main idea* (RI.3.2), selected-response items could be used to assess *determine the main idea of a text; recount the key details,* while a constructed-response item could be used to assess *explain how they support the main idea.* A computer-administered assessment, item response types include matching tables, fill-in tables, select or order text or graphics, and drag and drop.

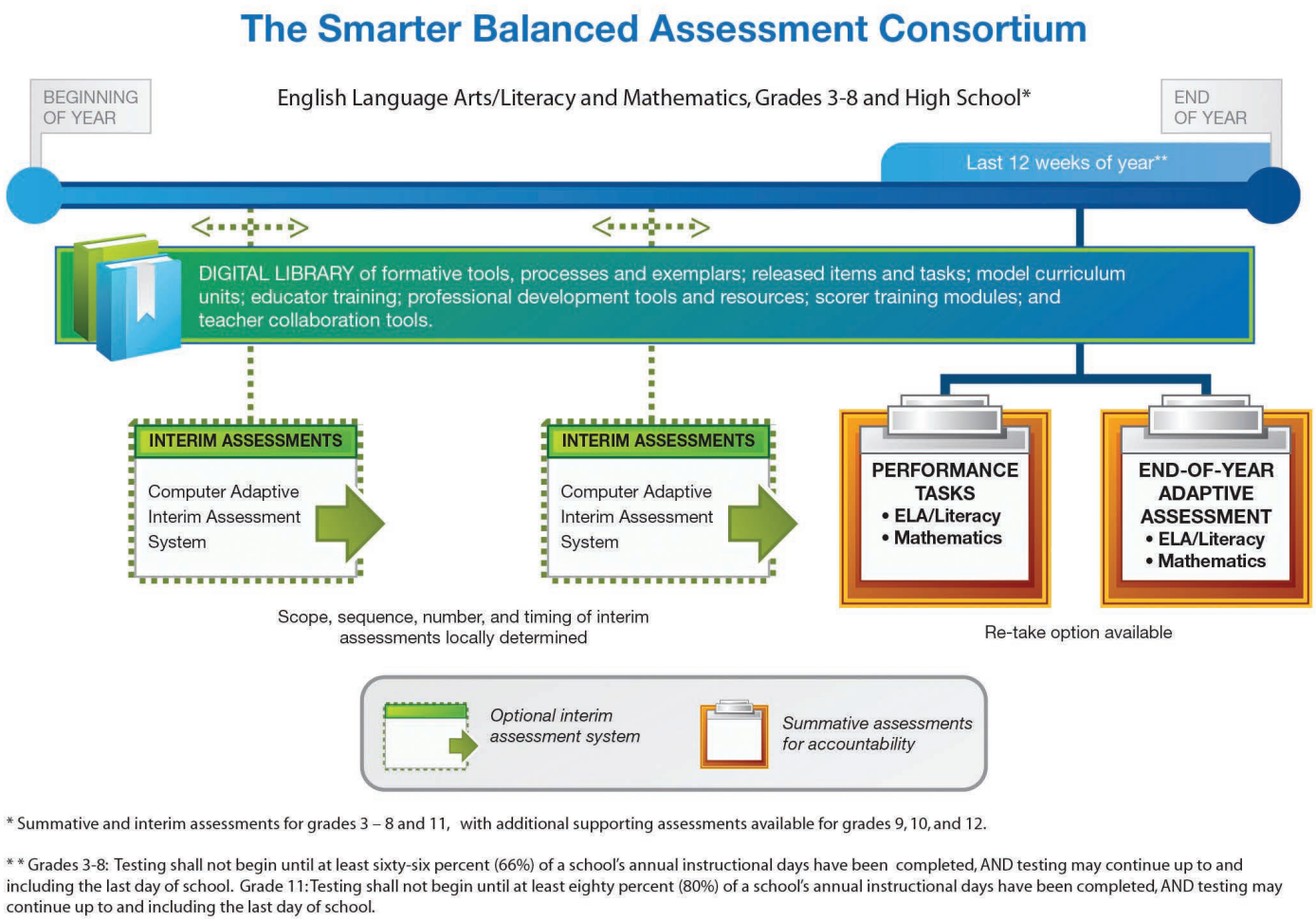
Locally determined measures are used to assess the achievement of students in kindergarten through second grade and to assess reading standards for foundational skills for kindergarten through grade five that are critical to every student's success in reading. The foundational skills are assessed intensively at kindergarten through grade two and then strategically at grade levels above grade two. In selecting appropriate assessments for these purposes, school district leaders need to refer to the section on the technical quality of assessments in this chapter to ensure that the assessments used are appropriate for their intended purposes.

*In addition, Smarter Balanced has a digital library of formative practices and tools for teachers' use. These tools include model units and lessons with embedded formative assessment strategies for teacher use.*

Optional interim assessments developed by the Smarter Balanced Assessment Consortium are also available to be administered at locally determined intervals. The interim assessments are reported on the same scale as the year-end assessments and permit teachers to assess either clusters of standards (referred to as Interim Assessment Blocks) or the full range of the CA CCSS for ELA/Literacy (referred to as Interim Comprehensive Assessments). In addition, Smarter Balanced has a digital library of formative practices and tools for teachers' use. These tools include model units and lessons with embedded formative assessment strategies for teacher use.

The Smarter Balanced Summative and Interim Assessments are computer adaptive tests and include performance tasks. These are described in more detail in figure 8.10.

***Figure 8.10. Smarter Balanced Assessment Consortium System***




* Summative and interim assessments for grades 3 – 8 and 11, with additional supporting assessments available for grades 9, 10, and 12.

** Grades 3-8: Testing shall not begin until at least sixty-six percent (66%) of a school's annual instructional days have been completed, AND testing may continue up to and including the last day of school. Grade 11: Testing shall not begin until at least eighty percent (80%) of a school's annual instructional days have been completed, AND testing may continue up to and including the last day of school.

**Source**

K–12 Center at Educational Testing Service. 2014. "The Smarter Balanced Assessment Consortium." Diagram created by Educational Testing Service.

## Computer Adaptive Tests

Computer-adaptive tests (CAT) tailor an assessment to individual students by presenting items based on a student's performance or responses to previous items in the test (Smarter Balanced 2013a). The Smarter Balanced Summative Assessments use CAT technology. The CAT assessment "engine" begins by delivering a short series of moderately difficult grade-level test items to the student, and then, depending on the student's initial performance, delivers items that are either more or less difficult. This process continues until the student's level of proficiency is determined (Smarter Balanced 2013a). For example, if a student has performed well on prior items, then more difficult items are given thereafter, but if a student has performed poorly on prior items, then easier items are presented. By matching the difficulty of new items more closely with a student's demonstrated level of performance, fewer items are needed. Some of the competencies assessed by CAT items include students' ability to use evidence to support their analyses (e.g., claims, conclusions, inferences) from reading different levels of text and their ability to edit and revise writing samples of different levels of complexity.

Because the test is administered by computer, it is critical that students develop the necessary technology skills, such as keyboarding, manipulating a mouse, and using pull-down menus, and that students have ample experience with the devices they are to use during the Smarter Balanced Summative Assessments.

## Performance Tasks

Performance tasks provide opportunities for students to demonstrate learning in ways that "emulate the context or conditions in which the intended knowledge and skills are actually applied" (AERA, APA, and NCME 1999, 137). They can take the form of demonstrations, oral performances, investigations, and written products (Lane 2013). Performance assessments provide better possibilities to measure complex skills and communication, important competencies, and disciplinary knowledge needed in today's society (Palm 2008) and important learning goals that cannot be easily assessed with other formats (Resnick and Resnick 1992).

*Performance assessments provide better possibilities to measure complex skills and communication, important competencies, and disciplinary knowledge needed in today's society and important learning goals that cannot be easily assessed with other formats.*

The performance tasks included in the Smarter Balanced Summative Assessments emphasize deep knowledge of core concepts and ideas, analysis, synthesis, communication, and critical thinking. They include several connected assessment items and may require more than one class period to complete. For example, to assess writing standards a performance task may ask students to write a full composition involving planning and revision in response to students' reading and analysis of multiple and varied texts. Similarly, performance tasks are used to assess grade 6–12 reading and writing standards for literacy in history/social studies, science, and technical subjects. For instance, short research projects that involve applying research and inquiry as well as a demonstration of many 21st century skills to produce a range of products (e.g., script for a presentation, multimedia presentation, public service announcement) are assessed with end-of-year performance tasks. Other constructed-response tasks include asking students to respond to a question about a passage they have read and use details from the text to support their answer, to write an ending to a story by adding details to tell what happens next, revising a paragraph by adding details to support an argument, and highlighting parts of a text that provide evidence to support a core idea of the text.

## Assessments for Students with Significant Cognitive Disabilities

The CCSS for ELA/Literacy are for every student, including students with significant cognitive disabilities. All students with disabilities participate in statewide assessments, with the exception of students who cannot achieve at or near grade level as identified by the members of the IEP team. These students present the most significant cognitive disabilities and make up approximately one percent of the population. They require substantial supports provided in connection with an alternative assessment. These supports allow identified students to have meaningful access to certain standards and assessment experiences as appropriate to their academic and functional needs. On October 1, 2012, California joined the National Center and State Collaborative (NCSC). The NCSC is currently developing several resources to support students with significant cognitive disabilities, including professional learning modules, curriculum models, instructional materials, alternate achievement standards, and a multi-state comprehensive assessment system. The long-term goal is to ensure that students with significant cognitive disabilities achieve increasingly higher academic outcomes and leave high school ready for postsecondary options. The NCSC is intended to be a standards-aligned assessment and is targeted to replace the previous alternate performance-based assessment known as the California Alternate Performance Assessment (CAPA). For more information, contact the California Department of Education Common Core Resources for Special Education Web site http://www.cde.ca.gov/sp/se/cc/.

*All students with disabilities participate in statewide assessments, with the exception of students who cannot achieve at or near grade level as identified by the members of the IEP team. These students present the most significant cognitive disabilities and make up approximately one percent of the population.*

## Biliteracy Assessment

When instruction is provided in English and in an additional language in alternative bilingual or dual language programs, classroom assessment for academic and language development progress in both languages is necessary. Such assessments should be designed according to the same principles and recommendations articulated throughout this *ELA/ELD Framework* and in this chapter both for ELs and for students whose primary language is English. Frequently and closely monitoring students' progress, assessing in both languages used for instruction, and interpreting assessment results in accordance with the research on effective bilingual education practices help ensure that students make steady and consistent progress toward full biliteracy and academic achievement in both languages. (English learners who have been enrolled in a U.S. school for less than 12 months do not to take the ELA portion of the Smarter Balanced Summative Assessments.)

*When instruction is provided in English and in an additional language in alternative bilingual or dual language programs, classroom assessment for academic and language development progress in both languages is necessary.*

## English Language Proficiency Assessments

The English Language Proficiency Assessments for California (ELPAC), which will be aligned to the CA ELD Standards adopted in 2012, is being developed to replace the California English Language Development Test (CELDT). The CELDT will be administered as usual until the ELPAC is fully operational.

The ELPAC will consist of two separate assessments: an initial assessment and a summative assessment. A summary of the identification and assessment process for ELs follows:

- **Home Language Survey (HLS):** School districts will continue to employ an HLS as the first step in identifying students whose primary language is not English. The HLS indicates if a student speaks a language other than English at home sometimes or all of the time. The HLS helps to determine which students are potentially EL and triggers the requirement to administer the ELPAC Initial Assessment to confirm EL classification.

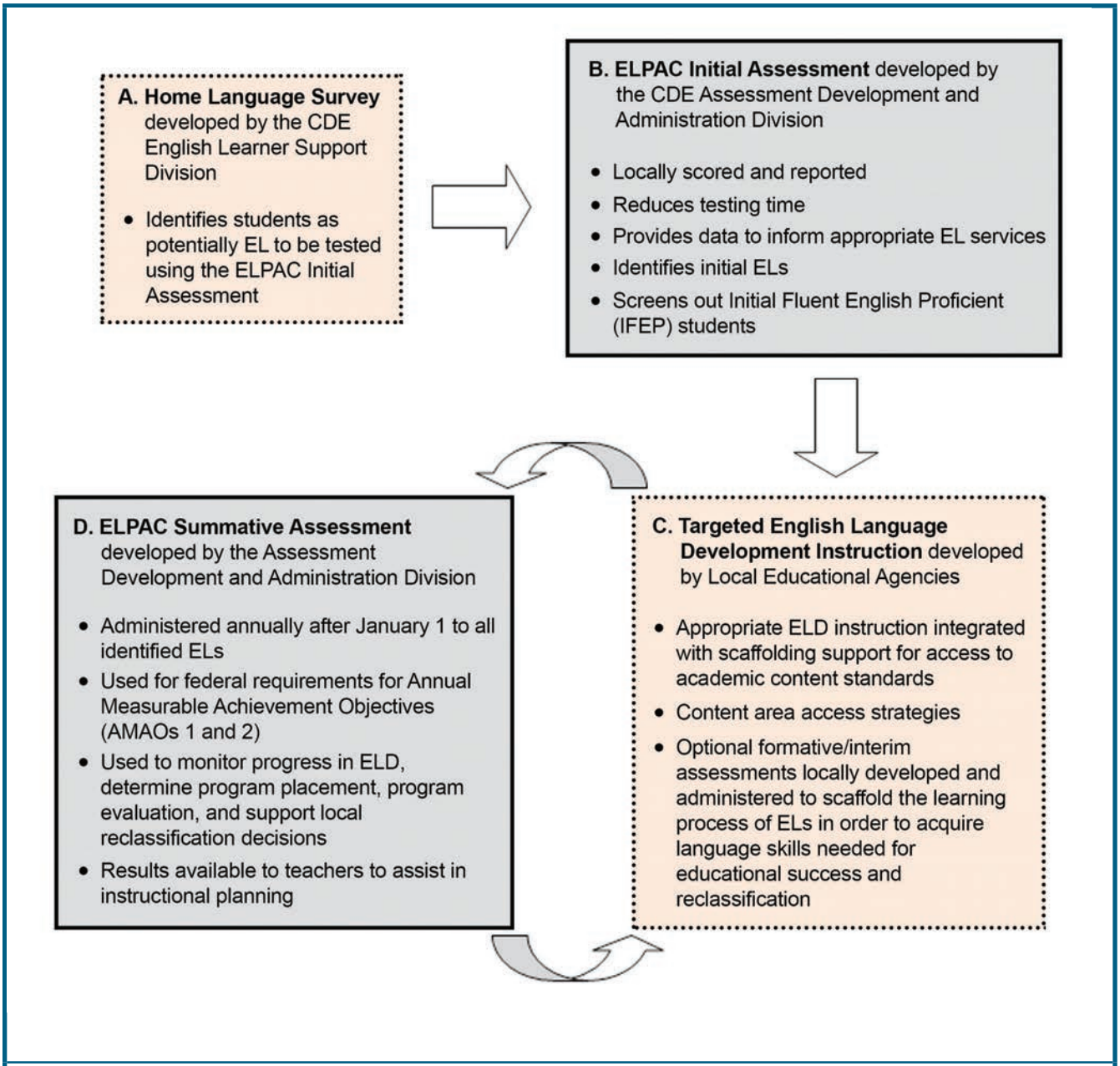
- **ELPAC Initial Assessment (Initial):** The Initial will be used by school districts to determine whether a student is an EL. It will be scored at the local level by qualified ELPAC examiners, resulting in a quicker turnaround of test results and timelier determination of EL classification and placement of students in appropriate instructional programs.

- **ELPAC Summative Assessment (Summative):** ELPAC examiners will annually administer the Summative to all identified ELs during a four-month period after January 1 as determined by the State Superintendent with the approval of the State Board of Education. The results will be used to determine ELs' annual progress in learning English for federal accountability purposes. The results may also be used by school districts to evaluate the effectiveness of their ELD programs, curricular resources, and instruction.

*The English Language Proficiency Assessments for California (ELPAC), which will be aligned to the CA ELD Standards adopted in 2012, is being developed to replace the California English Language Development Test (CELDT). The CELDT will be administered as usual until the ELPAC is fully operational.*

The ELPAC conceptual model (figure 8.11 below) highlights the process for using the HLS, the Initial, and the annual Summative. Boxes A and C have dotted borders to indicate local activities, and Boxes B and D have solid borders to indicate them as integral components of the state assessment system.

*Figure 8.11. ELPAC Conceptual Model*




**Source**

California Department of Education. 2013b. Memorandum "California English Language Development Test: Transitioning to the English Language Proficiency Assessment for California." Sacramento: California Department of Education.

# Technical Quality of Assessments

When considering the use of assessments to determine student achievement of the CA CCSS for ELA/Literacy and the CA ELD Standards, it is important to keep in mind the purpose for which a given assessment is intended. If an assessment does not permit proper inferences and provide accurate information for the specific decision-making purpose, its use may constitute misuse (Herman, Aschbacher, and Winters 1992).

This section elaborates the intended purpose of assessment. It is particularly important to refer to this section when selecting assessments other than California mandated assessments (e.g., Smarter Balanced Summative Assessments) whose technical quality are established through rigorous studies.

## Elements of Technical Quality

The idea of the *technical quality* of assessment refers the accuracy of information yielded by assessments and the appropriateness of the assessments for their intended purposes. There are three important elements related to the technical quality of assessments: validity, reliability, and freedom from bias (AERA, APA, and NCME 1999). Each element is described here, and figure 8.12, which summarizes the key points for each, is included at the end of this section.

### *Validity*

Validity is the overarching concept that defines quality in educational measurement. It is the extent to which an assessment permits appropriate inferences about student learning and contributes to the adequacy and appropriateness of using assessment results for specific decision-making purposes (Herman, Heritage, and Goldschmidt 2011). No assessment is valid for all purposes. While people often refer to the validity of a test, it is more correct to refer to the validity of the inferences or interpretations that can be made from the results of a test. Validity is basically a matter of degree; based on its purpose, an assessment can have high, moderate or low validity. For example, a diagnostic reading test might have a high degree of validity for identifying the type of decoding problems a student is having, a moderate degree for diagnosing comprehension problems, a low degree for identifying vocabulary knowledge difficulties, and no validity for diagnosing writing conventions difficulties. Similarly, annual assessments at the end of sixth grade have a high degree of validity for assessing achievement of standards for those students but no validity for assessing the achievement of the incoming group of sixth graders.

*No assessment is valid for all purposes. While people often refer to the validity of a test, it is more correct to refer to the validity of the inferences or interpretations that can be made from the results of a test.*

For an assessment to be valid for the intended purpose, there should be evidence that it does, in fact, assess what it purports to assess. Test publisher manuals should include information about the types of validity evidence that have been collected to support the intended uses specified for the assessment.

### *Reliability*

Reliability refers to how consistently an assessment measures what it is intended to measure (Linn and Miller 2005). If an assessment is reliable, the results should be replicable. For instance, changes in the time of administration, day and time of scoring, who scores the assessment, and the sample of assessment items should not create inconsistencies in results.

Reliability is important because it is a necessary adjunct of assessment validity (Linn and Miller 2005). If assessment results are not consistent, then it is reasonable to conclude that the results do not accurately measure what the assessment is purported to measure. A general rule of thumb for reliability is that the more items on an assessment the higher the reliability. Reliability is assessed primarily with

*Reliability refers to how consistently an assessment measures what it is intended to measure. If an assessment is reliable, the results should be replicable.*

statistical indices. Publishers' manuals should provide information about the reliability evidence for an assessment and the relevant statistical indices.

A variety of factors can influence the reliability of an assessment. For example, if a test is administered in an extremely hot or noisy room, students may not be able to complete the test to the best of their ability. If students are asked to provide an oral presentation when the instructions or expectations have not been made clear, this affects the reliability of the performance assessment. A number of other factors, including students' health, level of stress, and motivation can affect the reliability of an assessment. Teachers should use their judgment in interpreting assessment results when they suspect students are not able to perform to the best of their abilities. It is equally important for teachers to understand that a test or performance assessment may be reliable but not valid. For example, a student may consistently do well on an assessment, but the assessment may not be measuring what it claims to measure.

*Bias arises from tests that favor students of a particular gender, ethnicity, cultural background, geographic location, disability, or primary language. An assessment that is free from bias produces the same scores for students of the same attainment level, irrespective of their demographic subgroup.*

## Freedom from Bias

Bias can occur in test design or the way results are interpreted and used. Bias systematically disadvantages a student or group of students so that the students are unable to accurately show what they know and can do with respect to the content of the assessment. As a result, the assessment results may underestimate the students' achievement or reflect abilities that are not related to the assessment's content (Abedi and Lord 2001). Bias arises from tests that favor students of a particular gender, ethnicity, cultural background, geographic location, disability, or primary language. An assessment that is free from bias produces the same scores for students of the same attainment level, irrespective of their demographic subgroup.

Popham (1995) identifies two forms of bias: offensiveness and unfair penalization. Offensiveness occurs when the content of an assessment offends, upsets, or distresses particular subgroups, thus negatively influencing the test performance of these students. Items that present stereotypes of girls, boys, or particular cultures, or that portray certain groups as inferior, could adversely affect certain students' performance.

Unfair penalization occurs when the test content makes the test more difficult for some students than for others. Bias may occur, for example, if a test includes vocabulary that is unfamiliar to students because of their culture or geographic location. Bias may also occur if the test contains images that are more familiar to one group than another, or demands language skills beyond those of the targeted students. For example, if a reading assessment contains vocabulary related to rural life, then inner city students are potentially more disadvantaged than rural students. In addition, bias occurs when assessments that are based on letter-sound principles are used with students who do not have access to the sounds of language (i.e., students who are deaf or hard of hearing).

Assessment developers typically go to great lengths to make sure assessment items are not biased. Examine the publisher's manual for evidence that item reviews to guard against bias have been conducted.

Validity, reliability, and freedom from bias are all necessary conditions for assessment. They are not interchangeable (Linn and Miller 2005). For example, an assessment may offer consistent results (high reliability) without measuring what was targeted (low validity); and conversely a measurement with all the hallmarks of validity may not have high reliability. The key points of technical quality are summarized in figure 8.12.

***Figure 8.12. Key Points in Technical Quality of Assessments: Long- and Medium-Cycle Assessments***

| Technical Quality | Key Points |
|---|---|
| **Validity** | • Assessments need to be valid for the intended purpose<br>• The extent to which the information the assessment provides is accurate, adequate, and appropriate for a specific decision-making purpose<br>• While people often refer to the "validity of a test," it is more correct to refer to the validity of the *interpretations* that can be made from the results of a test<br>• No test is valid for all purposes |
| **Reliability** | • Consistency of the test results, repeatedly and over time<br>• Results of a test are reliable if they are replicable (despite changes in test administration and scoring, e.g., time of administration or who scores a test)<br>• Reliability is important because it is a necessary, but not sufficient, condition for validity. If assessment results are not consistent, then it is reasonable to conclude that the scores do not accurately measure what the test is intended to measure |
| **Freedom from Bias** | • Information or condition in an assessment that unfairly disadvantages a student or groups from showing their knowledge in the content<br>• An assessment free from bias produces same scores for students at the same attainment level, despite students' demographics (e.g., gender, ethnicity, primary language)<br>• Two forms of bias: (1) offensiveness – content offends or upsets particular subgroups, (2) unfair penalization – content more difficult for some students than others |

In the next section, the ideas of validity, reliability, and bias are considered in the context of formative assessment practice.

## Technical Quality and Formative Assessment

In formative assessment, the evidence generated by a variety of means is intended to provide information about students' learning progress in relation to specific learning goals (i.e., for a lesson) and to be used to inform immediate decisions about next steps in teaching and learning. Just as alignment to goals is important for annual and interim assessments, so it is for formative assessment. Teachers need to be clear about the specific learning goals (what students will learn, *not* what they will do) and what a successful performance entails. For example, learning goals for third-grade readers might be to (1) understand that the main idea is the author's message about a topic, minus all the details; and (2) determine the main idea of a text. The performances of understanding and skills for these goals would be for the students to (1) explain the main idea of a text, (2) locate where the author directly expresses the main idea (message) in text, and (3) explain how the important details describe the main idea. The teacher aligns her evidence gathering strategies with the goals and performance criteria.

For assessment to be formative it must be both timely and produce information that can inform teaching practice during its ongoing course (Erickson 2007). For this reason, the immediate or proximate timing of evidence is a key component of formative assessment validity. In addition, for formative assessment to be valid the resulting information must yield substantive insights into students' current learning status that can be used in subsequent pedagogical action (Heritage 2013).

*For assessment to be formative it must be both timely and produce information that can inform teaching practice during its ongoing course. For this reason, the immediate or proximate timing of evidence is a key component of formative assessment validity.*

An important point about validity in formative assessment concerns the consequences of assessment use. Because action resulting from the use of formative assessment evidence is intended to produce benefits to student learning, consequences represent an important component of the validity of such assessment. Even if assessments are formative in intention they are not so in practice if they do not generate further learning (Stobart 2006; Wiliam and Black 1996).

Reliability for classroom formative assessment takes a very different form because errors in instructional decisions can be rectified quickly by gathering more evidence of learning (Shepard 2001). Reliability in relation to instructional decisions can be thought of as "sufficiency of information" (Smith 2003, 30). In other words, teachers have to be confident that they have enough information about a student's learning to make a reasonable judgment about the current status of that learning. This idea of sufficiency of information for reliability argues for multiple sources of evidence before a teacher makes an instructional decision. The wider the range of information, and the more frequently the information is collected, the more accurately learning can be inferred (Griffin, and others 2010). In practical terms, this might mean that before making a judgment about student learning on specific features of language, a teacher has evidence from students' oral language production, from a quick-write, and from a text that has been underlined by students to identify the specific language feature in question. The more this kind of evidence is gathered in the context of everyday learning tasks, the less time is taken away from instruction and the more reliable the evidence gathered about a student's learning is (Linn and Baker 1996).

Because reading, writing, speaking, and listening skills do not develop in lockstep across all students, formative assessment is inevitably personalized and teachers need to employ strategies that tap into individual's knowledge and skills. Whatever evidence sources a teacher selects, they should account for the range of students present in the class so that all students have the opportunity to show where they are in their learning and have the prospect of moving forward from their current status. For example, well-designed questions and tasks that are sufficiently open-ended give all students the opportunity to reveal their learning. Similarly, formative assessment should not include any elements that would prevent some students from showing where they are relative to goals. These key points regarding technical quality in formative assessment are summarized in figure 8.13.

*. . . teachers have to be confident that they have enough information about a student's learning to make a reasonable judgment about the current status of that learning. This idea of sufficiency of information for reliability argues for multiple sources of evidence before a teacher makes an instructional decision.*

***Figure 8.13. Key Points in Technical Quality of Assessments: Short-Cycle Formative Assessment***

- Evidence gathered by the teacher is aligned to specific student learning goals derived from standards
- Evidence gathered is timely and contains information that can inform teaching
- Validity of formative assessment mainly lies in the use of evidence: information gathered yields substantive insights to students' current learning status that will be used for pedagogical action in order to move students toward achieving learning goals
- Reliability pertains to gathering enough information (e.g., multiple sources) about student learning in order to make a reasonable, accurate judgment for subsequent instructional decisions
- To ensure freedom from bias, evidence gathering is personalized to students so all students have the opportunity to show where they are in their learning and have the prospect of moving forward from their current learning status.

# Conclusion

Skilled use of assessment tools and processes is critical for ensuring students' achievement of the CA CCSS for ELA/Literacy and the CA ELD Standards. Only when teachers and leaders have a range of accurate information about student learning are they in a position to make decisions that advance learning. Key to informing the decisions educators need to make is a balanced and comprehensive system of assessment that provides different levels of detail for different decision-making purposes. Within such an assessment system, districts and school personnel need to strike the right balance in terms of the range of available assessments to teachers from the state or district, to those adopted by individual schools, to assessments embedded in curriculum materials, to ongoing day-by-day formative assessment practices that teachers engage in during instruction.

Assessment operates in the service of learning and involves careful consideration of the decisions that teachers need to make, when during the school year they need to make them to ensure student progress, and the assessment tools and processes they need to inform their decision-making. In combination with the right assessments for the right purposes, teachers' skillful use of assessment to support learning is critical to ensure that students in California meet the ambitious language and literacy standards that have been set forth.

# Works Cited


Abedi, Jamal. 2002. "Assessing and Accommodations of English Language Learners: Issues, Concerns and Recommendations." *Journal of School Improvement* 3 (1): 83–89.

Abedi, Jamal, and Nancy Ewers. 2013. *Accommodations for English Language Learners and Students with Disabilities: A Research-Based Decision Algorithm*. Smarter Balanced Assessment Consortium. http://www.smarterbalanced.org/wordpress/wp-content/uploads/2012/08/Accomodations-for-under-represented-students.pdf (accessed March 6, 2015). Link no longer available.

Abedi, Jamal, and Carol Lord. 2001. *Assessment and Accommodations for English Language Learners: Issues and Recommendations.* (CRESST Policy Brief No. 4). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing.

Allal, Linda. 2010. "Assessment and the Regulation of Learning." In *International Encyclopedia of Education,* Vol. 3, 3rd ed., edited by Penelope Peterson, Eva Baker, and Barry McGaw, 348–352. Oxford: Elsevier.

American Educational Research Association (AERA), American Psychological Association (APA), and National Council on Measurement in Education (NCME). 1999. *Standards for Educational and Psychological Testing*. Washington, DC: American Educational Research Association.

Ames, Carole, and Jennifer Archer. 1988. "Achievement Goals in the Classroom: Students' Learning Strategies and Motivation Processes." *Journal of Educational Psychology* 80: 260–267.

Andrade, Heidi G. 2013. *Essay Scoring Rubric.* Unpublished document.

Andrade, Heidi, Xiaolei Wang, Ying Du, and Robin L. Akawi. 2009. "Rubric-Referenced Self-Assessment and Self-Efficacy for Writing." *The Journal of Educational Research* 102 (4): 287–302.

Arter, Judith A., and Jan Chappuis. 2006. *Creating and Recognizing Quality Rubrics*. Portland, OR: Educational Testing Service.

Arter, Judith, and Vicki Spandel. 1992. "Using Portfolios of Student Work in Instruction and Assessment." *Educational Measurement: Issues and Practice* 11 (1): 36–44.

Bailey, Alison L., and Margaret Heritage. 2008. *Formative Assessment for Literacy, Grades K–6: Building Reading and Academic Language Skills Across the Curriculum*. Thousand Oaks, CA: Corwin/Sage Press.

Bangert-Drowns, Robert L., Chen-Lin C. Kulik, James A. Kulik, and Mary Teresa Morgan. 1991. "The Instructional Effect of Feedback in Test-like Events." *Review of Educational Research* 61: 213–238.

Bell, Beverley, and Bronwen Cowie. 2000. "The Characteristics of Formative Assessment in Science Education." *Science Education* 85: 536–553.

Black, Paul J., and Dylan Wiliam. 1998. "Assessment and Classroom Learning." *Assessment in Education: Principles, Policy, and Practice* 5 (1): 7–73.

Black, Paul, Christine Harrison, Clare Lee, Bethan Marshall, and Dylan Wiliam. 2003. *Assessment for Learning: Putting It into Practice*. New York, NY: Open University Press.

Boekaerts, Monique. 2006. Self-Regulation and Effort Investment. In *Handbook of Child Psychology:* Vol. 4. *Child Psychology in Practice,* 6th ed., edited by K. A. Renninger and I.E. Siegel, 345–377. New York: Wiley.

Bowles, Melissa, and Charles W. Stansfield. 2008. *A Practical Guide to Standards-Based Assessment in the Native Language* (NLA—LEP Partnership). http://www.ncela.us/files/rcd/BE024208/A_Practical_Guide.pdf (accessed June 10, 2013).

Brookhart, Susan M. 2013. *How to Create and Use Rubrics for Formative Assessment and Grading.* Alexandria, VA: ASCD.

Bunch, George C., Amanda Kibler, and Susan Pimental. 2012. "Realizing Opportunities for English Learners in the Common Core English Language Arts and Disciplinary Literacy Standards." In *Understanding Language: Commissioned Papers on Language and Literacy Issues in the Common Core State Standards and Next Generation Science Standards*, edited by Kenji Hakuta and M. Santos, 1–16. Palo Alto, CA: Stanford University.

Butler, Deborah L., and Philip H. Winne. 1995. "Feedback and Self-Regulated Learning: A Theoretical Synthesis." *Review of Educational Research* 65 (3): 245–281.

California Department of Education (CDE). 2013a. *California Common Core State Standards: English Language Arts and Literacy in History/Social Studies, Science, and Technical Subjects.* Sacramento: California Department of Education. http://www.cde.ca.gov/be/st/ss/documents/finalelaccssstandards.pdf (accessed April 5, 2013).

———. 2013b. Memorandum "California English Language Development Test: Transitioning to the English Language Proficiency Assessment for California." Sacramento: California Department of Education.

———. 2014 *California English Language Development Standards.* Sacramento: California Department of Education.

Carnegie Council on Advancing Adolescent Literacy. 2010. *Time to Act: An Agenda for Advancing Adolescent Literacy for College and Career Success.* New York: Carnegie Corporation of New York.

Chappuis, Jan, Rick J. Stiggins, Steve Chappuis, and Judith. A. Arter. 2012. *Classroom Assessment for Student Learning: Doing It Right – Using It Well.* 2nd ed. Portland, OR: Pearson Assessment Training Institute.

Christie, Frances, ed. 1999. *Pedagogy and the Shaping of Consciousness: Linguistic and Social Processes.* London, UK: Cassell Academic.

Cohen, Nicole, Tracey E. Hall, Ge Vue, and Patti Ganley. 2011. "Becoming Strategic Readers: Three Cases Using Formative Assessment, UDL, and Technology to Support Struggling Middle School Readers." In *New Frontiers in Formative Assessment*, edited by Pendred E. Noyce and Daniel T. Hickey, 129–140. Cambridge, MA: Harvard Education Press.

Council of Chief State School Officers (CCSSO) Formative Assessment State Collaborative. 2006. *Definition of Formative Assessment.* Washington, DC: Council of Chief State School Officers.

Crooks, Terence J. 1988. "The Impact of Classroom Evaluation on Students." *Review of Educational Research* 58: 438–481.

Derewianka, Beverly. 2011. *A New Grammar Companion for Teachers.* Sydney, NSW: Primary English Teaching Association.

Dweck, Carol S. 1999. *Self-Theories: Their Role in Motivation, Personality and Development.* Philadelphia, PA: Psychology Press.

———. 2006. *Mindset: The New Psychology of Success.* New York: Random House.

Erickson, Frederick. 2007. "Some Thoughts on 'Proximal' Formative Assessment of Student Learning." *Yearbook of the National Society for the Study of Education* 106: 186–216.

Fang, Zhihui and Zhujun Wang. 2011. "Beyond Rubrics: Using Functional Language Analysis to Evaluate Student Writing." *Australian Journal of Language and Literacy* 34 (2): 147–165.

Gersten, Russell, Donald Compton, Carol M. Connor, Joseph Dimino, Lana Santoro, Sylvia Linan-Thompson, and W. David Tilly. 2009. *Assisting Students Struggling with Reading Response to Intervention and Multi-Tier Intervention for Reading in the Primary Grades.* Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Gibbons, Pauline. 2009. *English Learners, Academic Literacy, and Thinking: Learning in the Challenge Zone.* Portsmouth, NH: Heinemann.

Griffin, Patrick. 2007. The Comfort of Competence and the Uncertainty of Assessment. *Studies in Educational Evaluation* 33: 87–99.

Griffin, Patrick, Leanne Murray, Esther Care, Amanda Thomas, and Pierina Perri. 2010. "Developmental Assessment: Lifting Literacy Through Professional Learning Teams." *Assessment in Education: Principles, Policy and Practice* 17 (4): 383–397.

Hacker, Douglas J., John Dunlosky, and Arthur C. Graesser, ed. 1998. *Metacognition in Educational Theory and Practice.* Mahwah, NJ: Lawrence Erlbaum.

Harlen, Wynne. 2007. "Formative Classroom Assessment in Science and Mathematics." In *Formative Classroom Assessment: Theory into Practice*, edited by J. H. McMillan, 116–135. New York: Teachers College Press.

Harlen, Wynne, and Ruth Deakin Crick. 2002. *A Systematic Review of the Impact of Summative Assessment and Tests on Student Motivation for Learning.* London: EPPI-Centre, Social Science Research Unit, Institute of Education. http://eppi.ioe.ac.uk/cms/LinkClick.aspx?fileticket=Pbyl1CdsDJU%3D&tabid=106&mid=1967 (accessed February 2013).

Harlen, Wynne, and Mary James. 1997. "Assessment and Learning: Differences and Relationships between Formative and Summative Assessment." *Assessment in Education: Principles, Policy and Practice* 4 (3): 365–379.

Hattie, John, and Helen Timperley. 2007. "The Power of Feedback." *Review of Educational Research* 77 (1): 81–112.

Heritage, Margaret. 2010. *Formative Assessment: Making It Happen in the Classroom*. Thousand Oaks, CA: Corwin Press.

———. 2013. *Formative Assessment in Practice: A Process of Inquiry and Action*. Cambridge, MA: Harvard Education Press.

———. 2014. Formative Assessment as the Key to Instructional Practice. Webinar provided by the Regional Educational Laboratory, Mid-Atlantic. June 19, 2014.

Herman, Joan L. 2010. *Coherence: Key to Next Generation Assessment Success*. Los Angeles, CA: University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST). http://www.cse.ucla.edu/products/policy/coherence_v6.pdf (accessed March 6, 2015).

Herman, Joan L., Pamela R. Aschbacher, and Lynn Winters. 1992. *A Practical Guide to Alternative Assessment*. Alexandria, VA: ASCD.

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.

Herman, Joan L., Margaret Heritage, and Pete Goldschmidt. 2011. *Developing and Selecting Assessments of Student Growth for Use in Teacher Evaluation Systems.* Los Angeles, CA: CRESST.

Herman, Joan L., Ellen Osmundson, and Ronald Dietel. 2010. *Benchmark Assessment for Improved Learning* (AACC Report). Los Angeles, CA: University of California.

International Reading Association. 2000. *Teaching All Children to Read: The Roles of the Reading Specialist: A Position Statement of the International Reading Association.* https://resources.oncourse.iu.edu/access/content/user/mikuleck/Filemanager_Public_Files/L650%20Specialist%20Roles/Roles-IRA%20position%20statement.pdf (accessed August 23, 2016).

Jenkins, Joseph R. 2003. *Candidate Measures for Screening At-Risk Students.* Paper presented at the National Research Center on Learning Disabilities Responsiveness-to-Intervention Symposium, Kansas City, MO.

Joseph, Laurice M. 2002. "Best Practices in Planning Intervention for Students with Reading Problems." In *Best Practices in School Psychology IV*, edited by Alex Thomas, and Jeff Grimes, 803–816. Bethesda, MD: National Association of School Psychologists.

K–12 Center at ETS. 2014. "The Smarter Balanced Assessment Consortium." Diagram created by Educational Testing Service. http://sas.sao.k12.hi.us/STATE/SAO/SASWebsite.nsf/10d1a575953d0e908a256c340001adab/226a4a6b992995e10a25796200068a3e/$FILE/SBAC-ETSChart.pdf (accessed August 23, 2016).

Kluger, Avraham N., and Angelo DeNisi. 1996. "The Effects of Feedback Interventions on Performance: A Historical Review, a Meta-Analysis, and a Preliminary Feedback Intervention Theory." *Psychological Bulletin* 119: 254–284.

Lane, Suzanne. 2013. "Performance Assessment." In *Sage Handbook of Research on Classroom Assessment*, edited by James H. McMillan, 313–331. Thousand Oaks, CA: SAGE Publications.

Linn, Robert L., and Eva L. Baker. 1996. "Can Performance-Based Student Assessment Be Psychometrically Sound?" In *Performance-Based Assessment: Challenges and Possibilities: Yearbook of the National Society for the Study of Education*, edited by Kenneth J. Rehage, Joan Boykoff Baron, and Dennie Palmer Wolf, 95 (1): 84-103). Chicago, IL: University of Chicago Press for the National Society for the Study of Education.

Linn, Robert L., and Michael David Miller. 2005. *Measurement and Assessment in Teaching*. Upper Saddle River, NJ: Pearson Education.

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. Washington, DC: Council of Chief State School Officers.

Lodico, Marguerite G., Elizabeth S. Ghatala, Joel R. Levin, Michael Pressley, and John A. Bell. 1983. "The Effects of Strategy-Monitoring Training on Children's Selection of Effective Memory Strategies." *Journal of Experimental Child Psychology* 35 (2): 263–277.

McManus, Sarah. 2008. *Attributes of Effective Formative Assessment.* Paper prepared for the Formative Assessment for Students and Teachers/State Collaborative on Assessment and Student Standards of the Council of Chief State School Officers. http://www.ccsso.org/Resources/Publications/Attributes_of_Effective_Formative_Assessment.html (accessed March 6, 2015).

National Association of State Directors of Special Education (NASDSE). 2005. *Response to Intervention: Policy Considerations and Implementation.* Alexandria, VA: Author.

National Research Council. 1999. *Improving Student Learning: A Strategic Plan for Education Research and its Utilization*. Washington, DC: The National Academy Press.

———. 2001. *Knowing What Students Know: The Science of Design and Educational Assessment*. Washington, DC: The National Academy Press.

———. 2012. *Education for Life and Work: Developing Transferable Knowledge and Skills in the 21st Century,* edited by James W. Pellegrino and Margaret L. Hilton. Committee on Defining Deeper Learning and 21st Century Skills, Board on Testing and Assessment and Board on Science Education, Division of Behavioral and Social Sciences Education. Washington, DC: The National Academy Press.

National Research Center on Learning Disabilities (NRCLD). 2006. *Progress Monitoring.* http://www.ldaofky.org/RTI/RTI%20Manual%20Section%202%20-%20Process%20Monitoring.pdf (accessed March 6, 2015).

Organisation for Economic Co-operation and Development (OECD). 2005. *Formative Assessment: Improving Student Learning in Secondary Classrooms*. Centre for Educational Research and Innovation. http://www.oecd.org/dataoecd/19/31/35661078.pdf (accessed May 9, 2013).

Palm, Torulf. 2008. "Performance Assessment and Authentic Assessment: A Conceptual Analysis of the Literature." *Practical Assessment, Research, and Evaluation* 13 (4): 1–11.

Pennock-Roman, Maria, and Charlene Rivera. 2011. "Mean Effects of Test Accommodations for ELLs and Non-ELLs: A Meta-Analysis of Experimental Studies." *Educational Measurement: Issues and Practice* 30 (3), 10–28.

Perie, Mariane, Scott Marion, and Brian Gong. 2009. "Moving Toward a Comprehensive Assessment System: A Framework for Considering Interim Assessments." *Educational Measurement: Issues and Practice* 28 (3): 5–13.

Perry, Nancy, Karen O. VandeKamp, Louise K. Mercer, and Carla J. Nordby. 2002. "Investigating Teacher-Student Interactions that Foster Self-Regulated Learning." *Educational Psychologist* 37 (1): 5–15.

Pintrich, Paul R. 2000. "The Role of Goal Orientation in Self-Regulated Learning." In *Handbook of Self-Regulation*, edited by Monique Boekaerts, Paul R. Pintrich, and Moshe Zeidner, 451–502. San Diego, CA: Academic Press.

Popham, W. James. 1995. *Classroom Assessment: What Teachers Need to Know.* Boston: Allyn and Bacon.

———. 2010. *Transformative Assessment.* Alexandria, VA: ASCD.

Poppers, Audrey E. 2011. "Identifying Craft Moves: Close Observation of Elementary Students' Writing." In *New Frontiers in Formative Assessment*, edited by Pendred E. Noyce and Daniel T. Hickey, 89–107. Cambridge, MA: Harvard Education Press.

Puckett, Margaret B., and Deborah Diffily. 2004. *Teaching Young Children: An Introduction to the Early Childhood Profession.* 2nd ed. Clifton Park, NY: Delmar Learning.

Resnick, Lauren B., and Daniel P. Resnick. 1992. "Assessing the Thinking Curriculum: New Tools for Educational Reform." In *Changing Assessments: Alternative Views of Aptitude, Achievement and Instruction*, edited by Bernard R. Gifford and Mary Catherine O'Connor, 37–76. New York: Kluwer/Springer Science+Business Media.

Ruiz-Primo, Maria A., and Erin M. Furtak. 2004. *Informal Assessment of Students' Understanding of Scientific Inquiry*. Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA.

———. 2006. "Informal Formative Assessment and Scientific Inquiry: Exploring Teachers' Practices and Student Learning." *Educational Assessment* 11 (3 and 4): 237–263.

———. 2007. "Exploring Teachers' Informal Formative Assessment Practices and Students' Understanding in the Context of Scientific Inquiry." *Journal of Educational Research in Science Teaching* 44 (1): 57–84.

Sadler, D. Royce. 1989. Formative Assessment and the Design of Instructional Strategies. *Instructional Science* 18: 119–144.

Schunk, Dale H., and Barry J. Zimmerman. 2008. *Motivation and Self-Regulated Learning: Theory, Research, and Applications*. Mahwah, NJ: Lawrence Erlbaum.

Shepard, Lorrie A. 2000. "The Role of Assessment in a Learning Culture." *Educational Researcher* 29 (7): 4–14.

———. 2001. The Role of Classroom Assessment in Teaching and Learning. In *Handbook of Research on Teaching,* 4th ed., edited by Virginia Richardson, 1066–1101. Washington, DC: American Educational Research Association.

———. 2005a. "Linking Formative Assessment to Scaffolding." *Educational Leadership* 63 (3): 66–71.

———. 2005b. *Will Commercialization Enable or Destroy Formative Assessment?* (Original title: Formative assessment: Caveat emptor). Paper presented at the ETS Invitational Conference 2005: The Future of Assessment, Shaping Teaching and Learning, New York.

Smarter Balanced Accessibility Consortium (SBAC). 2013a. http://www.smarterbalanced.org/ (accessed March 6, 2015).

———. 2013b. *Smarter Balanced Accessibility and Accommodations: Meeting the Needs of All Students.* http://www.smarterbalanced.org/wordpress/wp-content/uploads/2012/07/SmarterBalanced_Accessibility_Factsheet.pdf (accessed March 6, 2015).

———. 2014. *Smarter Balanced Assessment Consortium: Usability, Accessibility, and Accommodations Guidelines.* http://www.smarterbalanced.org/wordpress/wp-content/uploads/2014/08/SmarterBalanced_Guidelines.pdf (accessed March 6, 2015).

Smith, Jeffrey K. 2003. "Reconsidering Reliability in Classroom Assessment and Grading." *Educational Measurement: Issues and Practices* 22 (4): 26–34.

Soltero-Gonzalez, Lucinda, Kathy Escamilla, and Susan Hopewell. 2012. "Changing Teachers' Perceptions About the Writing Abilities of Emerging Bilingual Students: Towards a Holistic Bilingual Perspective on Writing Assessment." *International Journal of Bilingual Education and Bilingualism* 15 (1): 71–94.

Spycher, Pamela. 2007. "Academic Writing of English Learning Adolescents: Learning to Use 'Although.'" *Journal of Second Language Writing* 14 (4):238–254.

Spycher, Pamela, and Karin Linn-Nieves. 2014. Reconstructing, Deconstructing, and Constructing Complex Texts. In *The Common Core State Standards in English Language Arts/Literacy for English Language Learners: Grades K–5*, edited by Pamela Spycher. Alexandria, Virginia: TESOL Press.

Stansfield, Charles W., and Melissa Bowles. 2006. "Study 2: Test Translation and State Assessment Policies for English Language Learners." In *State Assessment Policy and Practice for English Language Learners: A National Perspective*, edited by Charlene Rivera and Eric Collum, 177–313. Mahwah, NJ: Lawrence Erlbaum/Routledge.

Stobart, Gordon. 2006. "The Validity of Formative Assessment." In *Assessment and Learning*, edited by John Gardner, 133–146. London, UK: Sage.

Understanding Language. 2013. *Instructional Unit: Persuasion Across Time and Space*. Palo Alto, CA: Stanford University. http://ell.stanford.edu/sites/default/files/ela_archives/understanding_language_materials_Jan2013.pdf (accessed March 6, 2015).

Venn, John J. 2000. *Assessing Students with Special Needs.* 2nd ed. Upper Saddle River, NJ: Merrill.

Wiliam, Dylan. 2006. *Does Assessment Hinder Learning?* Paper presented at the ETS Invitational Seminar at the Institute of Civil Engineers, London, UK.

———. 2007. "Keeping Learning on Track: Classroom Assessment and the Regulation of Learning." In *Second Handbook of Mathematics Teaching and Learning*, edited by Frank K. Lester Jr., 1053–1098. Greenwich, CT: Information Age Publishing.

———. 2011. "What is Assessment for Learning?" *Studies in Educational Evaluation* 37: 3–14.

Wiliam, Dylan, and Paul J. Black. 1996. "Meanings and Consequences: A Basis for Distinguishing Formative and Summative Functions of Assessment?" *British Educational Research Journal* 22 (5): 537–548.