Session 9: Assessment

Assessment of student learning is a complex and controversial topic, especially in an era of increased emphasis on standardized testing and school accountability. Yet assessment is an essential part of education, both to gain insight into student learning progress and (as is sometimes forgotten) into whether or not teaching strategies are appropriate and effective. Grant Wiggins, a leading commentator on issues in educational assessment, notes that the root of the word "assessment" derives from "to sit with." In its original sense, it speaks to connecting with the person being "assessed" in such a way as to gain a close understanding of what they know and understand.

Historically, testing has been used as a way of identifying "the best and brightest" students to move into a limited number of slots in secondary and post-secondary education. In an era when educational resources were very limited, this may have made sense. But today, the goals are to increase rather than limit the number of students who succeed and go on to higher education. With this change follows different needs and purposes for testing.

Currently educators use a variety of ways of assessing what students know. Some of these methods resemble traditional tests; others aren't always recognizable as such. All of them are referred to broadly as *assessment*. The kind of assessment that is used depends on what information is desired.

One of the current practices in education is to create a balance of different kinds of tests in order to test for a broader range of student knowledge and ability. This is sometimes referred to as having "multiple measures" to assess students. For instance, it's important to find out that children know their math facts; it's also important to find out that children know how to think critically and solve complex problems. Well-designed multiple choice and short answer questions can be helpful for assessing factual and some conceptual knowledge. Performance tasks, portfolios of work over time, and presentation of individual or group projects are useful for assessing complex reasoning and deeper student understanding. The kind of test or assessment method that is most appropriate depends on the kind of learning a teacher chooses to assess.

Wiggins and many other educators believe, for example, that "tests" should be administered in settings that enable the student to explain or clarify answers. More than this, they argue strongly for assessments that allow students to demonstrate the full range of their abilities in the performance of a real-world related task. Wiggins notes, "There is room for the quiet techie and the show-off prima donna in plays; there is room for the slow, heavy lineman and for the small, fleet pass receiver in football. In professional work, too, there is room for choice and style in tasks, topics, and methodologies. Why must all students be tested in the same way and at the same time? Why should speed of recall be so well rewarded and slow answering be so heavily penalized in conventional testing?"

Assessment is now (at best) woven into *all* stages of education: **before** new material is presented (to assess current knowledge), **during** the learning process (to monitor how well a student is grasping a concept), and **after** the lesson's been taught (to see how well the student understands the new material, and to assess the effectiveness of the teaching). This lets a teacher assess and modify

the learning experience in time to make a difference—rather than discover a problem just as it's time to move on to another subject.

The reading for this session, from Classroom Assessment and the National Science Education Standards, provides a research base for the importance of understanding and improving the type of assessment in the classroom that improves learning and pays particular attention to the notion of *formative assessment.*

For further thoughts on assessment, elaboration on the Wiggins references above, and reflections on thinking like an assessor, check out the following citation:

Wiggins, G. and McTighe, J. (2001) Understanding by Design. Upper Saddle River, New Jersey, Merrill Prentice Hall.

2
The Case for Strengthening Assessment in the Science Classroom

This chapter provides a rationale for this report: It presents a research base for the importance of understanding and improving the assessments that occur daily in classrooms that can directly influence learning. Teachers, teacher educators, administrators, and policy makers may find this chapter particularly relevant.

The goals for school science projected in the *Standards* represent a significant shift from traditional school practice. The document presents science as something that students actively do, rather than something that is done to or for them by teachers and texts. Science covers not only important facts but requires that objects and events be described carefully, that questions be asked about what is seen, that explanations of natural phenomena be constructed and tested, and that the resulting ideas be communicated to other people. It emphasizes the role of evidence in drawing conclusions. It involves making connections between students' current understandings of natural phenomena and the knowledge accepted and valued in the scientific community. Science also entails problem solving and decision making in the process of applying such knowledge to new situations and asking new questions. It is a way of knowing and thinking. If teachers can determine how well their students are meeting these new goals and students can learn how to gauge their progress, both can use this information to inform teaching and learning. By doing so, a vision for school science becomes a reality:

> The *Standards* present a vision of a scientifically literate populace. They outline what students need to know, understand, and be able to do to be scientifically literate at different grade levels. They describe an educational

> system in which all students demonstrate high levels of performance, in which teachers are empowered to make the decisions essential for effective learning, in which interlocking communities of teachers and students are focused on learning science, and in which supportive educational programs and systems nurture achievement. (NRC, 1999, p. 2)

Expectations conveyed in the *Standards* call for assessment to meet the full range of goals for science education. To this end, a broad view of assessment is proffered, for example:

> Ideas about assessments have undergone important changes in recent years. In the new view, assessment and learning are two sides of the same coin... When students engage in assessments, they should learn from those assessments. (NRC, 1999, pp. 5-6)

Assessments that resonate with a standards-based reform agenda reflect the complexity of science as a discipline of interconnected ideas and as a way of thinking about the world. Assessments must not be only summative in nature, that is, offering a cumulative summary of achievement level, usually at the end of a unit or after a topic has been covered. These summative assessments can serve multiple purposes: they help to inform placement decisions and to communicate a judgment about performance to interested parties, including parents and students. Assessment also must become an integral and essential part of daily classroom activity.

Research supports the value that the standards document places on **formative** assessment in enriching students' understanding of science. Black and Wiliam (1998a) define formative assessment as, "all those activities undertaken by teachers and by their students [that] provide information to be used as feedback to modify the teaching and learning activities in which they are engaged" (p. 7). They conducted a major review of more than 250 articles and books (Black & Wiliam, 1998b) that present research evidence on assessment from several countries. The main conclusion as a result of their study was as follows:

> Standards are raised only by changes that are put into direct effect by teachers and students in classrooms. There is a body of firm evidence that formative assessment is an essential feature of classroom work and that development of it can raise standards. We know of no other way of raising standards for which such a strong *prima facie* case can be made on the basis of evidence of such large learning gains. (p. 19)

Assessment becomes formative in nature only when either the teacher or the student **uses** that information to inform teaching and/or to influence learning. Therefore, data from summative assessments can be used in formative ways. In practice, it may be difficult to differentiate the two, as they often begin to blur. Summative

assessments may be given before the end of a unit, leaving time for modification in instruction and time for student revision. Moreover, teachers are constantly assessing to see where students are in respect to goals. If they keep careful records of some of these regular assessments, they can be used to inform instruction and communicate with parents and the student in a more "summative" way.

The Black and Wiliam review cites evidence that ongoing assessments by teachers, combined with appropriate feedback to students, can have powerful, positive effects on student learning and achievement. They also report that the learning gains from systematic attention to formative assessment are larger than most of those found for any other educational interventions. Although such findings provide impressive evidence of classroom practices that really work in improving student understanding, they also report that such practices are currently underdeveloped in most classrooms.

Black and Wiliam also offer a cautionary comment: Although many features essential to fruitful development of formative assessment now can be identified, there is no single, simple recipe that teachers could adopt and follow. Even though they differ from one another in many of the details, a variety of approaches to enhancing formative assessment turn out to be successful in improving learning. Why? Teachers differ and so do their students. The opportunities provided in various schools for engaging in an active science program differ as well.

Similarly many theories have been proposed as guides to improving learning, and several of them have practical implications. Many of the classroom-assessment examples in this volume emphasize the social and community aspects of learning. Learning is not viewed as solely an individual activity. For example, teachers and students develop shared understandings of standards for quality work and certain kinds of feedback extend and deepen these understandings. Other theories of learning highlight somewhat different aspects: meta-cognition (thinking about thinking) (Hacker, Dunlosky, & Graesser, 1998), "zones of proximal development" (deepening understanding by probing accessible knowledge for which one is prepared and ready) (Vygotsky, 1962); and the related idea that the teacher's role is to provide "scaffolding" for learning (an intellectual framework to which new ideas might be related) (Wood, Bruner, & Ross, 1976). There also are suggestive practices associated with the concept of "self-regulated" learning in which personal attributes and opportunities to learn intersect (Schunk & Zimmerman, 1998).

A FRAMEWORK FOR FORMATIVE ASSESSMENT

Notwithstanding the differing conceptions and views about learning,

all the best uses of formative assessment in the classroom seem to have a single common and straightforward underpinning that is fundamental to good educational practice. All these theories of learning—and several others—are consistent with the teacher helping the student operate within a framework of three guiding questions:

1. Where are you trying to go? (identify and communicate the learning and performance goals);
2. Where are you now? (assess, or help the student to self-assess, current levels of understanding);
3. How can you get there? (help the student with strategies and skills to reach the goal).

Almost all the theories that bear on formative assessment in the classroom can be related to this framework. The main point here, however, is that assessment (the second point) linked to action (the third point) results in formative assessment (Sadler, 1989).

Of course, it is not always necessary or even possible to follow these steps in a formal or sequential way when actually teaching. In the usual give-and-take of classroom life, opportunities arise unexpectedly to reexamine the goal of learning, to revisit a student's current understanding of a concept, and to revise the path toward the goal.

Science Content

This three-step guideline also leads to the centrality of a feature that underlies all three questions: the science content. The improvement of **science education** is the goal of the *Standards*, but much of what has been said so far in this chapter relates to any subject, not specifically science. Yet the formative assessment process cannot be implemented well in any field of study without serious attention to the nature and level of the subject matter to be taught. Which concepts are most important and for what reasons? Are some ideas particularly fruitful for laying the foundations for further learning? Are some of them more closely related than others to what the students already know, and thus can be presumed to be more accessible? In choosing a main content goal and the associated intermediate goals, which may be stepping-stones toward understanding that goal, the **science teacher** will be exercising subject expertise by: committing to certain **aims** for learning science, knowing the **science concepts** that best relate to that goal, and by **professional understanding** of the ways by which students may make progress in understanding the concepts and skills that lead to those goals. Furthermore, the choice of ways to assess student work similarly will be guided by personal **pedagogical** knowledge of those obstacles that are commonly

encountered by students in learning the particular science concepts that are chosen. These important responsibilities and daily decisions regarding curriculum and assessment underscore the importance of science teachers having a solid background and understanding of the science subject matter that they teach.

THE TEACHER'S ROLE

Much of the responsibility for implementing the science standards rests with classroom teachers. Assessment is no exception. The *Standards* recognize the importance of a teacher's ongoing assessments and indicate that classroom teachers are in the position to best use assessment in powerful ways for both formative and summative purposes, including improving classroom practice, planning curricula, developing self-directed learners, reporting student progress, and investigating their own teaching practices. Teachers' participation in classroom activities, hour after hour, day after day, positions them to gain information and insight into their students' understandings, actions, interests, intentions, and motivation that would be difficult to glean from tests (Darling-Hammond, 1994; Moss, 1994, 1996). Teachers need not only to interpret the assessment-generated information, they also must use the information to adapt their teaching repertoires to the needs of their students.

Feedback—Cognitive and Affective

The usefulness and effectiveness of formative assessment depend, in part, on the quality and saliency of the information gathered in the first place and the appropriateness and relevance of subsequent actions. The quality of the feedback rather than its existence or absence is the central point (Bangert-Downs, Kulik, Kulik, & Morgan, 1991; Sadler, 1989).

With regard to feedback, research makes the case for the use of descriptive, criterion-based feedback as opposed to numerical scoring or letter grades without clear criteria (Butler & Neuman, 1995; Cameron & Pierce, 1994; Kluger & deNisi, 1996). For example, in a study conducted by Butler (1987) with a random sampling of students, individuals completed an assessment task and then received one of three types of feedback: (a) tailored, written remarks addressing criteria they were aware of before taking the assessment, (b) grades derived from scoring of previous work, or (c) both grades and comments. Scores on two subsequent tasks increased most significantly for those who received detailed comments, while scores declined for those who received both comments and grades. For those assigned grades only, scores declined and then increased between the second and third tasks.

Butler's research is important to

consider in light of current research in attribution theory (Skaalvik, 1990; Vispoel & Austin, 1995). Research shows that feedback that emphasizes learning goals leads to greater learning gains than feedback that emphasizes self-esteem (Ames, 1992; Butler, 1988; Dweck, 1986). With respect to feedback emphasizing self-esteem, high-performing students often attribute their performance to effort and low-performing students attribute their performance to lack of ability (Butler & Newman, 1995; Cameron & Pierce, 1994; Kluger & deNisi, 1996). Students who repeatedly receive a grade of C– often believe that a C– is all that they are capable of achieving. Comments can take the focus from such attribution of success, or lack thereof, to the quality of the work at hand and areas where it can be strengthened. While grades can sometimes contribute to a competitive classroom environment where performance is normative (judged in relation to that of peers), comments that attend to specified criteria for higher quality work can help provide students with the feedback they need to develop their understanding and make progress.

Although letter grades are the most prevalent form of feedback, Stiggins (2001) reminds educators that grades serve as a way to convey a lot of information as a symbol for ease in communication. The symbol, or letter, can be only as meaningful as the definitions of achievement that underpin them and the quality of the assessment that produced them.

Design, Selection, and Participation

To use ongoing assessment to best facilitate student growth, the teacher plays a key role in choosing and organizing student tasks in ways that encourage them to speak and write so as to elicit their levels of understanding. Although almost any sample of student work can provide an occasion for a rich assessment discussion and can provide the teacher with assessment information, teachers also **plan** for opportunities for students to discuss and display their levels of understanding. They also create situations and allocate time for students to examine and discuss guidelines for high-quality work. These tasks are demanding ones and are discussed further in Chapters 3 and 5.

The *Standards* promote greater emphasis on teachers "continuously assessing student understanding," on "assessing rich, well-structured knowledge," on "assessing scientific understanding and reasoning," on "students engaged in ongoing assessment of their work and that of others," and on "teachers [becoming] involved in the development of external assessments" (p. 100).

The point of this last emphasis is significant: The *Standards* seek to

extend a teacher's influence beyond the classroom, or even the school, by advocating active teacher involvement in the **development and interpretation** of externally designed assessment used primarily for accountability purposes. Recent reforms in several state-assessment policies and practices in other countries provide models of how a teacher's assessments and participation can become more integral to the external assessments. This topic is discussed in Chapters 4, 5, and 6. Teacher involvement at this level is important to consider for many reasons, not the least of which is the centrality of inquiry in the vision of science education advanced in the *Standards*. Inquiry is difficult to assess in a one-time test. A teacher's position in the classroom allows for personal judgments of one's abilities over an extended investigation that cannot be matched by any feasible external testing procedure.

THE STUDENT'S ROLE

Student participation in the assessment process becomes essential if the *Standards* are to be actualized for all students. Specifically, **self**-assessment becomes crucial for feedback to be used effectively. Students are the ones who must ultimately take action to bridge the gap between where they are and where they are heading (Sadler, 1989). Brown (1994) stresses the strategic element of being aware of particular strengths and weaknesses: "Effective learners operate best when they have insight into their own strengths and weaknesses and access to their own repertoires of strategies for learning" (p. 9).

Research shows the potential learning gains from engaging students in peer- and self-assessment strategies (Covington, 1992; Darling-Hammond, Ancess, & Falk, 1995; White & Frederiksen, 1998; Wolf, Bixby, Glen, & Gardner, 1991). In a controlled experiment in two middle school science classrooms, White and Frederiksen (1998) demonstrated increases in student achievement in the class where discussion was structured to promote reflective peer- and self-assessment. The control group participated in general discussions of the curriculum content for the same amount of time but did not show the same increase in student achievement. Traditionally low-attaining students demonstrated the most notable improvement. This last point should not be overlooked. Supporting all students in their quest for high performance in science is an underlying principle of the science education standards. As this research indicates, assessment can be a critical means of reaching the goal.

Although the White and Frederiksen study demonstrates improved performance with regular student self-reflection, involving students in the assessment process

can serve other purposes as well. In a year-long teacher-researcher collaborative project in an elementary classroom, Rudd and Gunstone (1993) helped foster self-assessment skills through questionnaires, concept maps, and self-assessment maps. The researchers cite the following as evidence of an enhanced learning environment:

- the development of students' abilities to plan and think through their goals and skills;
- the creation of student awareness of the importance of evaluating their own work;
- the students' abilities to evaluate each other's self-assessment and in then providing constructive criticism; and
- the students' abilities to manage resources and time more effectively.

Assessment provides opportunities to discuss and develop a common understanding of what constitutes quality work. Students can have substantive conversations about what constitutes a good lab investigation, a salient scientific response, an appropriate use of evidence, or an effective presentation. Such discussions can be preliminary to the difficult challenge of trying to develop detailed assessment rubrics—tools that provide detailed descriptions and criteria for varying performance levels used to assess student work or responses—to help gauge quality work in each of these dimensions and to help guide the production of quality work. Participating in assessment can provide students with opportunities to reflect on what they are learning in order to make coherent connections within and between subject matters (Cole, Coffey, & Goldman, 1999; Resnick & Resnick, 1991; Wiggins, 1998). In the process of such deliberation, students often generate many of the salient educational goals themselves (Duschl & Gitomer, 1997). The process increases their commitment to achieving them (Covington, 1992). Furthermore, the ability to self-assess is essential for becoming a self-directed, lifelong learner (NRC, 1996), one of the aims set forth in the *Standards.*

THE SCHOOL'S ROLE

In-depth case studies conducted by Darling-Hammond and colleagues (1995) report how teachers and students in five schools used assessment to inform instruction and stimulate greater learning. Their work reinforces that assessment that makes learning central cannot be separated from other aspects of schooling. By focusing on schools where assessment occurs through "real-world" challenges that engage students in the assessment process, the studies provide examples of the role that observation, logs, portfolios, journal writing, and self- and peer-assessment, can play in facilitating powerful learning.

Portfolios (collections of student

work), regular self-reflection and peer assessment, assessment conversations, journals, projects, class discussions, performances, well-planned quizzes and tests—any combination of these assessment activities—can support improved science learning. In many classrooms, teachers are engaged in powerful teaching practices where assessment and learning work in concert toward creating a meaningful learning environment that benefits all students. However, another message that comes across in Darling-Hammond's case-study work is that many other teachers face school, district, or state policies that thwart attempts to move toward the vision of assessment and learning set forth in the *Standards*. Therefore, ensuring that assessment supports student learning requires support throughout the entire educational system. The system level is the topic of Chapter 6.

ASSESSMENT AND HIGH STANDARDS

A major impetus behind the standards movement is the expectation that **all** students are to achieve the high standards. To reach that goal, greater attention to classroom assessment that supports learning becomes particularly compelling, and teachers and researchers need to focus attention on how classroom assessment can be used as a means to this end. Assessment tools that calculate solely how well student achievement measures up to the standards, however reliable, will not suffice. Table 2-1 outlines the changes that are relevant to formative assessment as stated in the *Standards* (p. 100). Assessment also must serve as a vehicle for **improving** the quality of learning for every student. There is a clear and indivisible connection among assessment, curriculum, and teaching.

TABLE 2-1 Changing Emphases of Assessment

Less Emphasis On	More Emphasis On
Assessing what is easily measured	Assessing what is most highly valued
Assessing discrete knowledge	Assessing rich, well-structured knowledge
Assessing scientific knowledge	Assessing scientific understanding and reasoning
Assessing to learn what students do not know	Assessing to learn what students understand
End-of-term assessments by teachers	Students engaged in ongoing assessment of their work and that of others
Development of external assessments by measurement experts alone	Teachers involved in the development of external assessments

SOURCE: NRC. (1996).

Although the availability and intelligent use of curriculum and materials is essential, as Darling-Hammond (1994) suggests, "[e]fforts to raise standards of learning and performance must rest in part on strategies to transform assessment" (p. 6).

MULTIPLE PURPOSES OF ASSESSMENT

Although responsibility for assessment falls to the entire educational system, teachers and students are the primary designers, collectors, and users of assessment data in the direct service of learning. Recognizing the unique position of the classroom teacher, the science standards seek to recognize, legitimate, and extend the purview of the teacher in a range of assessment purposes and practices. In a comprehensive and coherent assessment system, teachers must accommodate the range of purposes that classroom assessment must serve—from self-reflection on practice, to monitoring achievement for individual students and assigning grades, to gauging levels of engagement, to reporting to parents, to making decisions about the placement of students. Black (1997) categorizes the **purposes** of assessment into those concerned with (a) support of learning; (b) certification, which includes reporting individual achievement, or grading, placement and promotion; and (c) accountability. Table 2-2 presents a visual overview that highlights the distinctions among the types, purposes, and locus of influence, as well as who takes on the primary roles and responsibilities with respect to the assessment.

Because different people are making judgments about students for different purposes, there are often serious areas of overlap that lead to ambiguities and tensions. Teachers, for example, must balance their roles

TABLE 2-2 Types, Purposes, and Roles and Responsibilities for Assessment

Type	Purpose	Roles and Responsibilities
Formative	Improve learning Inform instruction	Student and teacher
Summative	Grading Placement Promotion	Teachers and external tests
	Accountability	External tests (and teacher)

as facilitator and coach to promote learning along with their role as judge when they assign grades at the end of the term. External assessors, who prepare the standardized tests, serve primarily a summative function. It is important to keep in mind the different uses of assessment, the people who have the major responsibilities for them, and the intended audience, especially when considering the mechanisms employed to collect evidence and the inferences drawn from the collected data. For teachers especially, because they must engage in both formative and summative assessment practices, it is necessary to identify and attempt to mitigate the existing tensions. The challenge for classroom teachers becomes one of recognizing the range of factors that constitute assessment activity and taking full advantage of them to advance curricular, instructional, and learning goals. The challenge for the system becomes one of providing teachers and their students with the structures and necessary support to do so. These challenges are further elaborated in the chapters that follow.

KEY POINTS

- Research shows that regular and high-quality formative assessment can have a powerful, positive effect on student learning and achievement. The achievement gains associated with systematic attention to formative assessment are greater than most other educational interventions.
- The *Standards* indicate that classroom teachers are in the position to best use assessment in powerful ways for both formative and summative purposes, including improving classroom practice, planning curricula, developing self-directed learners, reporting student progress, and investigating their own teaching practices.
- There is no single recipe or blueprint that all teachers can successfully adopt and follow.
- Student participation in assessment becomes essential if high standards are to be actualized for all students.
- Assessment tools that calculate solely how well student achievement measures up to the standards, however reliable, will not suffice. Assessment must serve as a vehicle for improving the quality of learning for every student.
- Teachers need support from the larger system to realize and to take advantage of the possibilities of good assessment.