

Classroom Assessment: Minute by Minute, Day by Day

Siobhan Leahy, Christine Lyon, Marnie Thompson and Dylan William

In classrooms that use assessment to support learning, teachers continually adapt instruction to meet student needs.

There is intuitive appeal in using assessment to support instruction: assessment *for* learning rather than assessment *of* learning. We have to test our students for many reasons. Obviously, such testing should be useful in guiding teaching. Many schools formally test students at the end of a marking period—that is, every 6 to 10 weeks—but the information from such tests is hard to use, for two reasons.

First, only a small amount of testing time can be allotted to each standard or skill covered in the marking period. Consequently, the test is better for monitoring overall levels of achievement than for diagnosing specific weaknesses.

Second, the information arrives too late to be useful. We can use the results to make broad adjustments to curriculum, such as reteaching or spending more time on a unit, or identifying teachers who appear to be especially successful at teaching particular units. But if educators are serious about using assessment to improve instruction, then we need more fine-grained assessments, and we need to use the information they yield to modify instruction as we teach.

Changing Gears

What we need is a shift from *quality control* in learning to *quality assurance*. Traditional approaches to instruction and assessment involve teaching some given material, and then, at the end of teaching, working out who has and hasn't learned it—akin to a quality control approach in manufacturing. In contrast, assessment *for* learning involves adjusting teaching as needed while the learning is still taking place—a quality assurance approach. Quality assurance also involves a shift of attention from teaching to learning. The emphasis is on what the students are getting out of the process rather than on what teachers are putting into it, reminiscent of the old joke that schools are places where children go to watch teachers work.

In a classroom that uses assessment to support learning, the divide between instruction and assessment blurs. Everything students do—such as conversing in groups, completing seatwork, answering and asking questions, working on projects, handing in homework assignments, even sitting silently and looking confused—is a potential source of information about how much they understand. The teacher who consciously uses assessment to support learning takes in this information, analyzes it, and makes instructional decisions that address the understandings and misunderstandings that these assessments reveal. The amount of information can be overwhelming—one teacher likened it to “negotiating a swiftly flowing river”—so a key part of using assessment for learning is figuring out how to hone in on a manageable range of alternatives.

Research indicates that using assessment for learning improves student achievement. About seven years ago, Paul Black and one of us, Dylan William, found that students taught by teachers who used assessment for learning achieved in six or seven months what would otherwise have taken a year (1998). More important, these improvements appeared to be consistent across countries (including Canada, England, Israel, Portugal, and the United States), as well as across age brackets and content areas. We also found, after working with teachers in England, that these gains in achievement could be sustained over extended periods of time. The gains even held up when we measured student achievement with externally mandated standardized tests (see William, Lee, Harrison, & Black, 2004). Using this research and these ideas as a starting point, we and other colleagues at Educational Testing Service (ETS) have been working for the last two years with elementary, middle, and high school teachers in Arizona, Delaware, Maryland, Massachusetts, New Jersey, New Mexico, and Pennsylvania.

We have deepened our understanding of how assessment for learning can work in U.S. classrooms, and we have learned from teachers about the challenges of integrating assessment into classroom instruction.

Our Work with Teachers

In 2003 and 2004, we explored a number of ways of introducing teachers to the key ideas of assessment for learning. In one model, we held a three-day workshop during the summer in which we introduced teachers to the main ideas of assessment for learning and the research that shows that it works. We then shared specific techniques that teachers could use in their classrooms to bring assessment to life. During the subsequent school year, we met monthly with these teachers, both to learn from them what really worked in their classrooms and to offer suggestions about ways in which they might develop their practice. We also observed their classroom practices to gauge the extent to which they were implementing assessment-for-learning techniques and to determine the effects that these techniques were having on student learning. In other models, we spaced out the three days of the summer institute over several months (for example, one day in March, one in April, and one in May) so that teachers could try out some of the techniques in their classes between meetings.

As we expected, different teachers found different techniques useful; what worked for some did not work for others. This confirmed for us that there could be no one-size-fits-all package. However, we did find a set of five broad strategies to be equally powerful for teachers of all content areas and at all grade levels:

- Clarifying and sharing learning intentions and criteria for success.
- Engineering effective classroom discussions, questions, and learning tasks.
- Providing feedback that moves learners forward.
- Activating students as the owners of their own learning.
- Activating students as instructional resources for one another.

We think of these strategies as nonnegotiable in that they define the territory of assessment for learning. More important, we know from the research and from our work with teachers that these strategies are desirable things to do in any classroom.

However, the way in which a teacher might implement one of these strategies with a particular class or at a particular time requires careful thought. A self-assessment technique that works for students learning math in the middle grades may not work in a 2nd grade writing lesson. Moreover, what works for one 7th grade pre-algebra class may not work for the 7th grade pre-algebra class down the hall because of differences in the students or teachers.

Given this variability, it is important to offer teachers a range of techniques for each strategy, making them responsible for deciding which techniques they will use and allowing them time and freedom to customize these techniques to meet the needs of their students.

Teachers have tried out, adapted, and invented dozens of techniques, reporting on the results in meetings and interviews (to date, we have cataloged more than 50 techniques, and we expect the list to expand to more than 100 in the coming year). Many of these techniques require only subtle changes in practice, yet research on the underlying strategies suggests that they have a high “gearing”—meaning that these small changes in practice can leverage large gains in student learning (see Black & Wiliam, 1998; Wiliam, 2005). Further, the teaching practices that support these strategies are low-tech, low-cost, and usually feasible for individual teachers to implement. In this way, they differ dramatically from large-scale interventions, such as class size reduction or curriculum overhauls. We offer here a brief sampling of techniques for implementing each of the five assessment-for-learning strategies.

Clarify and Share Intentions and Criteria

Low achievement is often the result of students failing to understand what teachers require of them (Black & Wiliam, 1998). Many teachers address this issue by posting the state standard or learning objective in a prominent place at the start of the lesson, but such an approach is rarely successful because the standards are not written in student-friendly language.

Teachers in our various projects have explored many ways of making their learning objectives and their criteria for success transparent to students. One common method involves circulating work samples, such as lab reports, that a previous year's class completed, in view of prompting a discussion about quality. Students decide which reports are good and analyze what's good about the good ones and what's lacking in the weaker ones. Teachers have also found that by choosing the samples carefully, they can tune the task to the capabilities of the class. Initially, a teacher might choose four or five samples at very different quality levels to get students to focus on broad criteria for quality. As students grow more skilled, however, teachers can challenge them with a number of samples of similar quality to force the students to become more critical and reflective.

Engineer Effective Classroom Discussion

Many teachers spend a considerable proportion of their instructional time in whole-class discussion or question-and-answer sessions, but these sessions tend to rehearse existing knowledge rather than create new knowledge for students. Moreover, teachers generally listen for the "correct" answer instead of listening for what they can learn about the students' thinking; as Davis (1997) says, they listen *evaluatively* rather than *interpretively*. The teachers with whom we have worked have tried to address this issue by asking students questions that either prompt students to think or provide teachers with information that they can use to adjust instruction to meet learning needs.

As a result of this focus, teachers have become aware of the need to carefully plan the questions that they use in class. Many of our teachers now spend more time planning instruction than grading student work, a practice that emphasizes the shift from quality control to quality assurance. By thinking more carefully about the questions they ask in class, teachers can check on students' understanding while the students are still in the class rather than after they have left, as is the case with grading. Some questions are designed as "range-finding" questions to reveal what students know at the beginning of an instructional sequence. For example, a high school biology teacher might ask the class how much water taken up by the roots of a corn plant is lost through transpiration. Many students believe that transpiration is "bad" and that plants try to minimize the amount of water lost in this process, whereas, in fact, the "lost" water plays an important role in transporting nutrients around the plant.

A middle school mathematics teacher might ask students to indicate how many fractions they can find between $\frac{1}{6}$ and $\frac{1}{7}$. Some students will think there aren't any; others may suggest an answer that, although in some way understandable, is an incorrect use of mathematical notation, such as 1 over $6\frac{1}{2}$. The important feature of such range-finding items is that they can help a teacher judge where to begin instruction.

Of course, teachers can use the same item in a number of ways, depending on the context. They could use the question about fractions at the end of a sequence of instruction on equivalent fractions to see whether students have grasped the main idea. A middle school science teacher might ask students at the end of a laboratory experiment, "What was the dependent variable in today's lab?" A social studies teacher, at the end of a project on World War II, might ask students to state their views about which year the war began and give reasons supporting their choice.

Teachers can also use questions to check on student understanding before continuing the lesson. We call this a "hinge point" in the lesson because the lesson can go in different directions, depending on student responses. By explicitly integrating these hinge points into instruction, teachers can make their teaching more responsive to their students' needs in real time.

However, no matter how good the hinge-point question, the traditional model of classroom questioning presents two additional problems. The first is lack of engagement. If the classroom rule dictates that students raise their hands to answer questions, then students can disengage from the classroom by keeping their hands down. For this reason, many of our teachers have instituted the idea of "no hands up, except to ask a question." The teacher can either decide whom to call on to answer a question or use some randomizing device, such as a beaker of Popsicle sticks with the students' names written on them. This way, all students know that they need to stay engaged because the teacher could call on any one of them. One teacher we worked with reported that her students love the fairness of this

approach and that her shyer students are showing greater confidence as a result of being invited to participate in this way. Other teachers have said that some students think it's unfair that they don't get a chance to show off when they know the answer.

The second problem with traditional questioning is that the teacher gets to hear only one student's thinking. To gauge the understanding of the whole class, the teacher needs to get responses from all the students in real time. One way to do this is to have all students write their answers on individual dry-erase boards, which they hold up at the teacher's request. The teacher can then scan responses for novel solutions as well as misconceptions. This technique would be particularly helpful with the fraction question we cited.

Another approach is to give each student a set of four cards labeled *A*, *B*, *C*, and *D*, and ask the question in multiple-choice format. If the question is well designed, the teacher can quickly judge the different levels of understanding in the class. If all students answer correctly, the teacher can move on. If no one answers correctly, the teacher might choose to reteach the concept. If some students answer correctly and some answer incorrectly, the teacher can use that knowledge to engineer a whole-class discussion on the concept or match up the students for peer teaching. Hinge-point questions provide a window into students' thinking and, at the same time, give the teacher some ideas about how to take the students' learning forward.

Provide Feedback That Moves Learners Forward

After the lesson, of course, comes grading. The problem with giving a student a grade and a supportive comment is that these practices don't cause further learning. Before they began thinking about assessment for learning, none of the teachers with whom we worked believed that their students spent as long considering teacher feedback as it had taken the teachers to provide that feedback. Indeed, the research shows that when students receive a grade and a comment, they ignore the comment (see Butler, 1988). The first thing they look at is the grade, and the second thing they look at is their neighbor's grade.

To be effective, feedback needs to cause thinking. Grades don't do that. Scores don't do that. And comments like "Good job" don't do that either. What *does* cause thinking is a comment that addresses what the student needs to do to improve, linked to rubrics where appropriate. Of course, it's difficult to give insightful comments when the assignment asked for 20 calculations or 20 historical dates, but even in these cases, feedback can cause thinking. For example, one approach that many of our teachers have found productive is to say to a student, "Five of these 20 answers are incorrect. Find them and fix them!"

Some of our teachers worried about the extra time needed to provide useful feedback. But once students engaged in self-assessment and peer assessment, the teachers were able to be more selective about which elements of student work they looked at, and they could focus on giving feedback that peers were unable to provide.

Teachers also worried about the reactions of administrators and parents. Some teachers needed waivers from principals to vary school policy (for example, to give comments rather than grades on interim assessments). Most principals were happy to permit these changes once teachers explained their reasons. Parents were also supportive. Some even said they found comments more useful than grades because the comments provided them with guidance on how to help their children.

Activate Students as Owners of Their Learning

Developing assessment for learning in one's classroom involves altering the implicit contract between teacher and students by creating shared responsibility for learning. One simple technique is to distribute green and red "traffic light" cards, which students "flash" to indicate their level of understanding (green = understand, red = don't understand). A teacher who uses this technique with her 9th grade algebra classes told us that one day she moved on too quickly, without scanning the students' cards. A student picked up her own card as well as her neighbors' cards, waved them in the air, and pointed at

them wildly, with the red side facing the teacher. The teacher considered this ample proof that this student was taking ownership of her learning.

Students also take ownership of their learning when they assess their own work, using agreed-on criteria for success. Teachers can provide students with a rubric written in student-friendly language, or the class can develop the rubric with the teacher's guidance (for examples, see Black, Harrison, Lee, Marshall, & Wiliam, 2003). The teachers we have worked with report that students' self-assessments are generally accurate, and students say that assessing their own work helped them understand the material in a new way.

Activate Students as Instructional Resources for One Another

Getting students started with self-assessment can be challenging. Many teachers provide students with rubrics but find that the students seem unable to use the rubrics to focus and improve their work. For many students, using a rubric to assess their own work is just too difficult. But as most teachers know, students from kindergarten to 12th grade are much better at spotting errors in other students' work than in their own work. For that reason, peer assessment and feedback can be an important part of effective instruction. Students who get feedback are not the only beneficiaries. Students who give feedback also benefit, sometimes more than the recipients. As they assess the work of a peer, they are forced to engage in understanding the rubric, but in the context of someone else's work, which is less emotionally charged. Also, students often communicate more effectively with one another than the teacher does, and the recipients of the feedback tend to be more engaged when the feedback comes from a peer. When the teacher gives feedback, students often just "sit there and take it" until the ordeal is over.

Using peer and self-assessment techniques frees up teacher time to plan better instruction or work more intensively with small groups of students. It's also a highly effective teaching strategy. One cautionary note is in order, however. In our view, students should not be giving another student a grade that will be reported to parents or administrators. Peer assessment should be focused on improvement, not on grading.

Using Evidence of Learning to Adapt Instruction

One final strategy binds the others together: Assessment information should be used to adapt instruction to meet student needs.

As teachers listen to student responses to a hinge-point question or note the prevalence of red or green cards, they can make on-the-fly decisions to review material or to pair up those who understand the concept with those who don't for some peer tutoring. Using the evidence they have elicited, teachers can make instructional decisions that they otherwise could not have made.

At the end of the lesson, many of the teachers with whom we work use "exit passes." Students are given index cards and must turn in their responses to a question posed by the teacher before they can leave the classroom. Sometimes this will be a "big idea" question, to check on the students' grasp of the content of the lesson. At other times, it will be a range-finding question, to help the teacher judge where to begin the next day's instruction.

Teachers using assessment for learning continually look for ways in which they can generate evidence of student learning, and they use this evidence to adapt their instruction to better meet their students' learning needs. They share the responsibility for learning with the learners; students know that they are responsible for alerting the teacher when they do not understand. Teachers design their instruction to yield evidence about student achievement; for example, they carefully craft hinge-point questions to create "moments of contingency," in which the direction of the instruction will depend on student responses. Teachers provide feedback that engages students, make time in class for students to work on improvement, and activate students as instructional resources for one another.

All this sounds like a lot of work, but according to our teachers, it doesn't take any more time than the practices they used to engage in. And these techniques are far more effective. Teachers tell us that they are enjoying their teaching more.

Supporting Teacher Change

None of these ideas is new, and a large and growing research base shows that implementing them yields substantial improvement in student learning. So why are these strategies and techniques not practiced more widely? The answer is that knowing about these techniques and strategies is one thing; figuring out how to make them work in your own classroom is something else.

That's why we're currently developing a set of tools and workshops to support teachers in developing a deep and practical understanding of assessment for learning, primarily through the vehicle of school-based teacher learning communities. After we introduce teachers to the basic principles of assessment for learning, we encourage them to try out two or three techniques in their own classrooms and to meet with other colleagues regularly—ideally every month—to discuss their experiences and see what the other teachers are doing (see Black, Harrison, Lee, Marshall, & Wiliam, 2003, 2004). Teachers are accountable because they know they will have to share their experiences with their colleagues. However, each teacher is also in control of what he or she tries out. Over time, the teacher learning community develops a shared language that enables teachers to talk to one another about what they are doing. Teachers build individual and collective skill and confidence in assessment for learning. Colleagues help them decide when it is time to move on to the next challenge as well as point out potential pitfalls.

In many ways, the teacher learning community approach is similar to the larger assessment-for-learning approach. Both focus on where learners are now, where they want to go, and how we can help them get there.

References

Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2003). *Assessment for learning: Putting it into practice*. Buckingham, UK: Open University Press.

Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2004). Working inside the black box: Assessment for learning in the classroom. *Phi Delta Kappan*, 86(1), 8–21.

Black, P., & Wiliam, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, 80(2), 139–147.

Butler, R. (1988). Enhancing and undermining intrinsic motivation. *British Journal of Educational Psychology*, 58, 1–14.

Davis, B. (1997). Listening for differences: An evolving conception of mathematics teaching. *Journal for Research in Mathematics Education*, 28(3), 355–376.

Wiliam, D. (2005). Keeping learning on track: Formative assessment and the regulation of learning. In M. Coupland, J. Anderson, & T. Spencer (Eds.), *Making mathematics vital: Proceedings of the twentieth biennial conference of the Australian Association of Mathematics Teachers* (pp. 26–40). Adelaide, Australia: Australian Association of Mathematics Teachers.

Wiliam, D., Lee, C., Harrison, C., & Black, P. J. (2004). Teachers developing assessment for learning: Impact on student achievement. *Assessment in Education: Principles, Policy & Practice*, 11(1), 49–65.

Siobhan Leahy, Christine Lyon, Marnie Thompson, and Dylan Wiliam (dwiliam@ets.org) work in the Learning and Teaching Research Center, Educational Testing Service, Rosedale Rd., Princeton, NJ 08541.