

Inside the Black Box: Raising Standards Through Classroom Assessment

Paul Black and Dylan Wiliam
Kappan, October 1998

In this highly influential article in the October 1998 issue of *Kappan*, British professors Paul Black and Dylan Wiliam discuss the evidence from 250 research studies on formative assessment and argue that when teachers use minute-by-minute, day-by-day assessments to check for understanding – and follow up by adapting instruction to clear up their students’ confusions and misunderstandings – they are on the road to getting high achievement from all their students. In fact, argue Black and Wiliam, this is a better way of raising standards than external programs, and is the most powerful engine for closing the achievement gap.

Black and Wiliam don’t think most school reform efforts pay enough attention to what goes on inside the classroom, treating it as a “black box” and assuming that the pressure of high-stakes tests will be enough to get teachers to improve learning. “[I]t seems strange, even unfair,” they argue, “to leave the most difficult piece of the standards-raising puzzle entirely to teachers. If there are ways in which policy makers and others can give direct help and support to the everyday classroom task of achieving better learning, then surely these ways ought to be pursued vigorously.” Black and Wiliam are on a mission to understand what goes on inside the “black box” and provide research support for teachers’ best efforts. Here are their major points:

• ***Formative assessments are a highly effective way to raise student achievement.*** Black and Wiliam say the research supports this conclusion for students from kindergarten to college, in different subject areas, and in different countries. Effect sizes range from 0.4 to 0.7, which is higher than most other educational interventions. Interestingly, formative assessments have their most positive effect on low-achieving students and therefore narrow the achievement gap.

But the authors admit that it’s not a simple matter to implement classroom assessments well. Research indicates that effective use of assessments depends on teachers making significant changes in classroom practice, modifying instruction based on formative data, getting students actively involved in the process, and using these assessments – and students’ self-assessments – to improve motivation and self-esteem.

• ***Classroom assessments are often problematic.*** Teachers have a strong tendency to teach, test, and move on with the curriculum – without using classroom data to adjust instruction or help struggling students. “Marking is usually conscientious but often fails to offer guidance on how work can be improved,” said a U.K. inspection report. “Information about pupil performance received by the teacher is insufficiently used to inform subsequent work.” There are several other problems with the way many teachers use classroom assessments:

- Many classroom tests measure rote and superficial learning even when teachers are trying to develop understanding; often teachers seem unaware of the inconsistency.
- Teachers seldom share their classroom assessments with their colleagues; without critical scrutiny, many of these tests are not high-quality.

- “For primary teachers particularly,” say Black and Wiliam, “there is a tendency to emphasize quantity and presentation of work and to neglect its quality in relation to learning.”
- Some teachers use tests for social and managerial purposes, rather than to diagnose students’ learning needs or give useful advice; when this happens, the impact on low-achieving students can be quite negative.
- Some teachers use classroom assessments to foster competition and compare students to one another; this can lead low-achieving students to believe that they lack “ability” and can’t learn.
- Some teachers use classroom tests to predict students’ performance on external tests, rather than to diagnose their students’ learning needs.
- Some teachers don’t pay attention to the assessment records of students’ previous teachers.

Black and Wiliam then list a number of ways that schools can improve the use of formative assessments:

• **Assessment feedback to students should be about the particular qualities of their work, accompanied by advice on how to improve it, and should not compare students to each other.** “When the classroom culture focuses on rewards, ‘gold stars,’ grades, or class ranking,” write Black and Wiliam, “then pupils look for ways to obtain the best marks rather than to improve their learning.” This leads students to avoid difficult tasks for fear of not being successful and losing points. Students who don’t do well come to believe that they lack innate ability and can’t do anything about that. They tend to avoid investing effort in learning, are content to “get over,” and build up their self-esteem in other arenas.

Effective use of formative assessments reverses this negative culture. The message to students is that their failures are temporary and can be fixed by effective effort. Classrooms that use assessments in this way have a culture of success, backed by a belief that all students can achieve. This dynamic is especially helpful for low-achieving students because the focus is on the specific problems they are having with their work and what they can do to improve.

• **Self-assessment by students is an essential component of formative assessment.** “When anyone is trying to learn,” say Black and Wiliam, “feedback about effort has three elements: recognition of the *desired goal*, evidence about *present position*, and some understanding of a *way to close the gap* between the two. All three must be understood to some degree by anyone before he or she can take action to improve learning.” First and foremost, students need to have a clear picture of what they are supposed to be learning. “Surprisingly and sadly,” continue the authors, “many pupils do not have such a picture, and they appear to have become accustomed to receiving classroom teaching as an arbitrary sequence of exercises with no overarching rationale. To overcome this pattern of passive reception requires hard and sustained work. When pupils do acquire such an overview, they then become more committed and more effective as learners.”

• **Teachers need to build in opportunities for students to communicate their evolving understanding as a unit unfolds – and give them adequate wait-time to respond.** Effective teachers plan open-ended questions, classroom tasks, and homework assignments that develop and assess student understanding of the big ideas and details of the unit in real time. Master teachers are also alert to unexpected student responses. Less effective teachers tend to listen for the “right” answer and, quite unconsciously, respond in ways that signal that there’s only one answer and the students’ job is to figure it out

and give it back. “Over time,” say Black and Wiliam of this pattern, “the pupils get the message: they are not required to think out their own answers. The object of the exercise is to work out – or guess – what answer the teacher expects to see or hear.”

When teachers ask questions and don’t give students enough time to think through their answers (providing answers to their own questions or moving quickly on to another student), two things happen: the questions tend to become low-level, with simple answers; and the classroom “dialogue” tends to take place between the teacher and a few students who are good at giving rapid-fire answers to a succession of questions. “So the teacher,” say Black and Wiliam, “by lowering the level of questions and by accepting answers from a few, can keep the lesson going but is actually out of touch with the understanding of most of the class. The question/answer dialogue becomes a ritual, one in which thoughtful involvement suffers.” Black and Wiliam suggest a number of ways to break out of this pattern:

- Giving students time to think after asking a question;
- Asking students to discuss their thinking in pairs or small groups, so that a respondent is speaking on behalf of others;
- Giving students a choice between different possible answers and asking them to vote on the options;
- Asking all students to write down an answer and then reading out a selected few.

“What is essential,” say Black and Wiliam, “is that any dialogue should evoke thoughtful reflection in which all pupils can be encouraged to take part, for only then can the formative process start to work.”

• ***Teachers need to get the most out of classroom tests.*** Tests should be opportunities for learning, insist Black and Wiliam. “It is better to have frequent short tests than infrequent long ones. Any new learning should first be tested within about a week of a first encounter...” Generating good test items that are relevant to the main learning aims and communicate clearly with students is difficult, say Black and Wiliam – which is a strong reason for teachers to collaborate and draw on outside sources.

Equally important is the quality of feedback that students get after a test. “Research studies have shown that, if pupils are given only marks or grades, they do not benefit from the feedback,” say Black and Wiliam. “Feedback has been shown to improve learning when it gives each pupil specific guidance on strengths and weaknesses, preferably without any overall marks... Pupils must be given the means and opportunities to work with evidence of their difficulties.”

“The worst scenario,” say Black and Wiliam, “is one in which some pupils who get low marks this time also got low marks last time and come to expect to get low marks next time. This cycle of repeated failure becomes part of a shared belief between such students and their teacher.”

• ***It’s quite challenging for teachers to implement formative assessments well.*** Some students will resist this change from the familiar classroom routine. Being asked to think for oneself and take responsibility for self-assessment is difficult. In addition, using formative assessments takes more classroom time at first. It’s a leap of faith for many teachers to make these changes – but it will produce wider and deeper learning in the long run.

Teachers need to deal with two other issues if they are serious about using formative assessments. The first is their theory of action about how students learn. If a teacher has bought into the transmission model of teaching, namely...

- Knowledge is transmitted by the teacher and learned by the student;
- Understanding will develop later;
- Clarity of exposition accompanied by rewards for patient reception are the essentials of good teaching

... then that teacher will not see the point of formative assessments. Fortunately most teachers have moved past the transmission model and know that teacher-student interaction is vital for learning to take place. These teachers should be receptive to improving their continuous assessment of student learning.

A second possible barrier arises when teachers have doubts about the potential of all their students to learn. There are two polarities:

- The “fixed intelligence” view – a belief that each student has a fixed, inherited intelligence that cannot be altered much by schooling;
- The “untapped potential” view – a belief that starts with the assumption that so-called ability is a complex of skills that can be learned.

Formative assessments are most useful for teachers who hold the second belief – who operate on the assumption that “all pupils can learn more effectively if one can clear away, by sensitive handling, the obstacles to learning, be they cognitive failures never diagnosed or damage to personal confidence or a combination of the two.” The “untapped potential” approach helps all students learn better, and has its most powerful impact on students who have previously struggled in school.

At the end of their article, Black and Wiliam express hope that these ideas on formative assessment will spread by a small number of schools piloting and developing them, then reach more schools through dissemination of best practices, with principals and teachers working to reduce the obstacles that now exist in many schools. “The chief negative influence here,” they say, “is that of short external tests. Such tests can dominate teachers’ work, and, insofar as they encourage drilling to produce right answers to short, out-of-context questions, they can lead teachers to act against their own better judgment about the best ways to develop the learning of their pupils.”

The key, say Black and Wiliam, is for teachers to be clear about the distinction between formative and summative tests and treat formative assessments as a low-stakes learning experience for their students and themselves. This, they conclude, is the most powerful way to raise the achievement of all students – especially disadvantaged students – on high-stakes summative tests and close the achievement gap.

“Inside the Black Box: Raising Standards Through Classroom Assessment” by Paul Black and Dylan Wiliam in *Phi Delta Kappan*, October 1998 (Vol. 80 #2, p. 139-148), <http://www.pdkintl.org/kappan/kbla9810.htm>

Summary from: Marshall Memo 146

A Weekly Round-up of Important Ideas and Research in K-12 Education, July 24, 2006