NCTM NATIONAL COUNCIL OF TEACHERS OF MATHEMATICS

Assessment
Research Brief

Five "Key Strategies" for Effective Formative Assessment

In order to build a comprehensive framework for formative assessment, Wiliam and Thompson (2007) proposed that three processes were central:

1. Establishing where learners are in their learning
2. Establishing where they are going
3. Establishing how to get there

By considering separately the roles of the teacher and the students themselves, they proposed that formative assessment could be built up from five "key strategies."

1. Clarifying, sharing, and understanding goals for learning and criteria for success with learners

There are a number of ways teachers can begin the process of clarifying and sharing learning goals and success criteria. Many teachers specify the learning goals for the lesson at the beginning of the lesson, but in doing so, many teachers fail to distinguish between the learning goals and the activities that will lead to the required learning. *When teachers start from what it is they want students to know and design their instruction backward from that goal, then instruction is far more likely to be effective (Wiggins and McTighe 2000).*

Wiggins and McTighe also advocate a two-stage process of first clarifying the learning goals themselves (what is worthy and requiring understanding?), which is then followed by establishing success criteria (what would count as evidence of understanding?). Only then should the teacher move on to exploring activities that will lead to the required understanding.

However, it is important that students also come to understand these goals and success criteria, as Royce Sadler (1989, p. 121) notes:

> The indispensable conditions for improvement are that the student comes to hold a concept of quality roughly similar to that held by the teacher, is continuously able to monitor the quality of what is being produced during the act of production itself, and has a repertoire of alternative moves or strategies from which to draw at any given point.

Indeed, there is evidence that discrepancies in beliefs about what it is that counts as learning in mathematics classrooms may be a significant factor in the achievement gaps observed in mathematics classrooms. In a study of 72 students between the ages of seven and thirteen, Gray and Tall (1994) found that the reasoning of the higher-achieving students was qualitatively different from that of the lower-achieving students. In particular, the higher-achieving students were able to work successfully despite unresolved ambiguities about whether mathematical entities were concepts or procedures. Lower-achieving students were unable to accept such ambiguities and could not work past them. By refusing to accept the ambiguities inherent in mathematics, the lower-achieving students were, in fact, attempting a far more difficult form of mathematics, with a far greater cognitive demand.

A simple example may be illustrative here. When we write $6\frac{1}{2}$, the mathematical operation between the 6 and the $\frac{1}{2}$ is actually addition, but when we write $6x$, the implied operation between the 6 and the x is multiplication, and the relationship between the 6 and the 1 in 61 is different again. And yet, very few people who are successful in mathematics are aware of these inconsistencies or differences in mathematical notation. In a very real sense, being successful in mathematics requires knowing what to worry about and what not to worry about. *Students who do not understand what is important and what is not important will be at a very real disadvantage.*

In a study of twelve seventh-grade science classrooms, White and Frederiksen (1998) found that giving students time to talk about what would count as quality work, and how their work was likely to be evaluated, reduced the achievement gap between the highest- and lowest-achieving students in half and increased the average performance of the classes to such an extent that the weakest students in the experimental group were outperforming all but the very strongest students in the control group.

This is why using a variety of examples of students' work from other classes can be extremely powerful in helping students come to understand what counts as quality work. Many teachers have found that students are better at spotting errors in the work of other students than they are at seeing them in their own work. By giving students examples of work at different standards, students can begin to explore the differences between superior and inferior work, and these emergent understandings can be discussed with the whole class.

As a result of such processes, students will develop a "nose for quality" (Claxton 1995) that they will then be able to use in monitoring the quality of their own work.

2. Engineering effective classroom discussions, questions, activities, and tasks that elicit evidence of students' learning

Once we know what it is that we want our students to learn, then it is important to collect the right sort of evidence about the extent of students' progress toward these goals, but few teachers plan the kinds of tasks, activities, and questions that they use with their students specifically to elicit the right kind of evidence of students' learning. As an example, consider the question shown in figure 1 below.

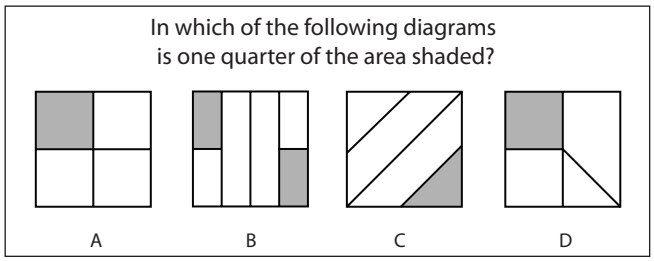

Fig. 1. Diagnostic item on elementary fractions

Diagram A is the obvious answer, but B is also correct. However, some students do not believe that one-quarter of B is shaded because of a belief that the shaded parts have to be adjoining. Students who believe that one-quarter of C is shaded have not understood that one region shaded out of four is not necessarily a quarter. Diagram D is perhaps the most interesting here. One-quarter of this diagram is shaded, although the pieces are not all equal; students who rely too literally on the "equal areas" definition of fractions will say that D is not a correct response. *By crafting questions that explicitly build in the undergeneralizations and overgeneralizations that students are known to make (Bransford, Brown, and Cocking 2000), we can get far more useful information about what to do next.* Furthermore, by equipping each student in the class with a set of four cards bearing the letters A, B, C, and D and by requiring all students to respond simultaneously with their answers, the teacher can generate a very solid evidence base for deciding whether the class is ready to move on (Leahy et al. 2005). If every student responds with A, B, and D, then the teacher can move on with confidence that the students have understood. If everyone simply responds with A, then the teacher may choose to reteach some part of the topic. The most likely response, however, is for some students to respond correctly and for others to respond incorrectly, or incompletely. This provides the teacher with an opportunity to conduct a classroom discussion in which students with different views can be asked to justify their selections.

Of course planning such questions takes time, but by investing the time before the lesson, the teacher is able to address students' confusion during the lesson, with the students still in front of him or her. Teachers who do not plan such questions are forced to put children's thinking back on track through grading, thus dealing with the students one at a time, after they have gone away.

3. Providing feedback that moves learning forward

The research on feedback shows that much of the feedback that students receive has, at best, no impact on learning and can actually be counterproductive. Kluger and DeNisi (1996) reviewed more than three thousand research reports on the effects of feedback in schools, colleges, and workplaces and found that only 131 studies were scientifically rigorous. In 50 of these studies, feedback actually made people's performance worse than it would have been without feedback. The principal feature of these studies was that feedback was, in the psychological jargon, "ego-involving." In other words, the feedback focused attention on the person rather than on the quality of the work——for example, by giving scores, grades, or other forms of report that encouraged comparison with others. The studies where feedback was most effective were those in which the feedback told participants not just what to do to improve but also how to go about it.

Given the emphasis on grading in U.S. schools, teachers may be tempted to offer comments alongside scores or grades. However, a number of studies (e.g., Butler 1987, 1988) have shown that when comments are accompanied by grades or scores, students focus first on their own grade or score and then on those of their neighbors, so that grades with comments are no more effective than grades alone, and much less effective than comments alone. The crucial requirement of feedback is that it should force the student to engage cognitively in the work.

Such feedback could be given orally, as in this example from Saphier (2005, p. 92):

> *Teacher:* What part don't you understand?
>
> *Student:* I just don't get it.
>
> *Teacher:* Well, the first part is just like the last problem you did. Then we add one more variable. See if you can find out what it is, and I'll come back in a few minutes.

Written feedback can support students in finding errors for themselves:

- There are 5 answers here that are incorrect. Find them and fix them.

- The answer to this question is … Can you find a way to work it out?

It can also identify where students might use and extend their existing knowledge:

- You've used substitution to solve all these simultaneous equations. Can you use elimination?

Other approaches (Hodgen and Wiliam 2006) include encouraging pupils to reflect:

- You used two different methods to solve these problems. What are the advantages and disadvantages of each?
- You have understood … well. Can you make up your own more difficult problems?

Another suggestion is to have students discuss their ideas with others:

- You seem to be confusing sine and cosine. Talk to Katie about how to work out the difference.
- Compare your work with Ali and write some advice to another student tackling this topic for the first time.

The important point in all this is that as well as "putting the ball back in the students' court," the teacher also needs to set aside time for students to read, respond to, and act on feedback.

4. Activating students as owners of their own learning

When teachers are told they are responsible for making sure that their students do well, the quality of their teaching deteriorates, as does their students' learning (Deci et al. 1982). In contrast, when students take an active part in monitoring and regulating their learning, then the rate of their learning is dramatically increased. Indeed, it is common to find studies in which the rate of students' learning is doubled, so that students learn in six months what students in control groups take a year to learn (Fontana and Fernandes 1994; Mevarech and Kramarski 1997).

In an attempt to integrate research on motivation, metacognition, self-esteem, self-efficacy, and attribution theory, Monique Boekaerts has proposed a dual-processing theory of student motivation and engagement (Boekaerts 2006). When presented with a task, the student evaluates the task according to its interest, difficulty, cost of engagement, and so on. If the evaluation is positive, the student is likely to seek to increase competence by engaging in the task. If the evaluation is negative, a range of possible outcomes is possible. The student may engage in the task but focus on getting a good grade from the teacher instead of mastering the relevant material (e.g., by cheating) or the student may disengage from the task on the grounds that "it is better to be thought lazy than dumb." *The important point for teachers is that to maximize learning, the focus needs to be on personal growth rather than on a comparison with others.*

Practical techniques for getting students started include "traffic lights," where students flash green, yellow, or red cards to indicate their level of understanding of a concept. Many teachers have reported that initially, students who are focusing on well-being, rather than growth, display green, indicating full understanding, even though they know they are confused. However, when the teacher asks students who have shown green cards to explain concepts to those who have shown yellow or red, students have a strong incentive to be honest!

5. Activating students as learning resources for one another

Slavin, Hurley, and Chamberlain (2003) have shown that activating students as learning resources for one another produces some of the largest gains seen in any educational interventions, provided two conditions are met. The first is that the learning environment must provide for group goals, so that students are working as a group instead of just working in a group. The second condition is individual accountability, so that each student is responsible for his or her contribution to the group, so there can be no "passengers."

With regard to assessment, then, a crucial feature is that the assessment encourages collaboration among students while they are learning. To achieve this collaboration, the learning goals and success criteria must be accessible to the students (see above), and the teacher must support the students as they learn how to help one another improve their work. One particularly successful format for doing this has been the idea of "two stars and a wish." The idea is that when students are commenting on the work of one another, they do not give evaluative feedback but instead have to identify two positive features of the work (two "stars") and one feature that they believe merits further attention (the "wish"). Teachers who have used this technique with students as young as five years old have been astonished to see how appropriate the comments are, and because the feedback comes from a peer rather than someone in authority over them, the recipient of the feedback appears to be more able to accept the feedback (in other words, they focus on growth rather than on preserving their well-being). In fact, teachers have told us that the feedback that students give to one another, although accurate, is far more hard-hitting and direct than they themselves would

have given. Furthermore, the research shows that the person providing the feedback benefits just as much as the recipient because she or he is forced to internalize the learning intentions and success criteria in the context of someone else's work, which is less emotionally charged than doing it in the context of one's own work.

Conclusion

The available research evidence suggests that considerable enhancements in student achievement are possible when teachers use assessment, minute-by-minute and day-by-day, to adjust their instruction to meet their students' learning needs. However, it is also clear that making such changes is much more than just adding a few routines to one's normal practice. It involves a change of focus from what the teacher is putting into the process and to what the learner is getting out of it, and the radical nature of the changes means that the support of colleagues is essential. Nevertheless, our experiences to date suggest that the investment of effort in these changes is amply rewarded. Students are more engaged in class, achieve higher standards, and teachers find their work more professionally fulfilling. As one teacher said, "I'm not babysitting any more."

By Dylan Wiliam
Judith Reed, Series Editor

REFERENCES

Boekaerts, Monique. "Self-Regulation and Effort Investment." In *Handbook of Child Psychology,* Vol. 4: *Child Psychology in Practice,* 6th ed., edited by K. Ann Renninger and Irving E. Sigel, pp. 345–77). Hoboken, N.J.: John Wiley & Sons, 2006.

Bransford, John D., Ann L. Brown, and Rodney R. Cocking. *How People Learn: Brain, Mind, Experience, and School.* Washington, D.C.: National Academies Press, 2000.

Butler, Ruth. "Task-Involving and Ego-Involving Properties of Evaluation: Effects of Different Feedback Conditions on Motivational Perceptions, Interest and Performance." *Journal of Educational Psychology* 79, no. 4 (1987): 474–82.

———. "Enhancing and Undermining Intrinsic Motivation: The Effects of Task-Involving and Ego-Involving Evaluation on Interest and Performance." *British Journal of Educational Psychology* 58 (1988): 1–14.

Claxton, G. L. "What Kind of Learning Does Self-Assessment Drive? Developing a 'Nose' for Quality: Comments on Klenowski." *Assessment in Education: Principles, Policy and Practice* 2, no. 3 (1995): 339–43.

Deci, Edward L., N. H. Speigel, R. M. Ryan, R. Koestner, and M. Kauffman. "The Effects of Performance Standards on Teaching Styles: The Behavior of Controlling Teachers." *Journal of Educational Psychology* 74 (1982): 852–59.

Fontana, David., and M. Fernandes. "Improvements in Mathematics Performance as a Consequence of Self-Assessment in Portuguese Primary School Pupils." *British Journal of Educational Psychology* 64, no. 4 (1994): 407–17.

Gray, Eddie M., and David O. Tall. "Duality, Ambiguity, and Flexibility: A 'Proceptual' View of Simple Arithmetic." *Journal for Research in Mathematics Education* 25 (March 1994): 116–40.

Hodgen, Jeremy, and Dylan Wiliam. *Mathematics inside the Black Box: Assessment for Learning in the Mathematics Classroom.* London: NFER-Nelson, 2006.

Kluger, Avraham N., and Angelo DeNisi. "The Effects of Feedback Interventions on Performance: A Historical Review, a Meta-analysis, and a Preliminary Feedback Intervention Theory." *Psychological Bulletin* 119, no. 2 (1996): 254–84.

Leahy, Siobhan, Christine Lyon, Marnie Thompson, and Dylan Wiliam. (2005). "Classroom Assessment: Minute-by-Minute and Day-by-Day." *Educational Leadership* 63, no. 3 (2005): 18–24.

Mevarech, Zemira R.., and Bracha Kramarski. "IMPROVE: A Multidimensional Method for Teaching Mathematics in Heterogeneous Classrooms." *American Educational Research Journal* 34, no. 2 (1997): 365–94.

Sadler, D. Royce. "Formative Assessment and the Design of Instructional Systems." *Instructional Science* 18, no. 2 (1989): 119–44.

Saphier, Jonathon. "Masters of Motivation." In *On Common Ground: The Power of Professional Learning Communities,* edited by Richard DuFour, Robert Eaker, and Rebecca DuFour, pp. 85–113. Bloomington, Ill.: National Education Service, 2005.

Slavin, Robert E., Eric A. Hurley, and Anne M. Chamberlain. "Cooperative Learning and Achievement." In *Handbook of Psychology,* Vol. 7: *Educational Psychology,* edited by W. M. Reynolds and G. J. Miller, pp. 177–98. Hoboken, N.J.: John Wiley & Sons, 2003.

White, Barbara Y., and John R. Frederiksen. "Inquiry, Modeling, and Metacognition: Making Science Accessible to All Students." *Cognition and Instruction* 16, no. 1 (1998): 3–118.

Wiggins, Grant, and Jay McTighe. *Understanding by Design.* New York: Prentice Hall, 2000.

Wiliam, Dylan, and Marnie Thompson. "Integrating Assessment with Instruction: What Will It Take to Make It Work?" In *The Future of Assessment: Shaping Teaching and Learning,* edited by C. A. Dwyer. Mahwah, N.J.: Lawrence Erlbaum Associates, 2007.