DOCUMENT RESUME

ED 480 068 CG 032 641

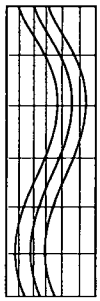
AUTHOR	Arter, Judith A.
TITLE	Assessment for Learning: Classroom Assessment To Improve Student Achievement and Well-Being.
PUB DATE	2003-08-00
NOTE	24p.; In: Measuring Up: Assessment Issues for Teachers, Counselors, and Administrators; see CG 032 608.
PUB TYPE	Information Analyses (070)
EDRS PRICE	EDRS Price MF01/PC01 Plus Postage.
DESCRIPTORS	*Academic Achievement; *Classroom Environment; *Educational Assessment; Evaluation Methods; Learning Strategies; *Well Being

ABSTRACT

This chapter describes classroom assessment for learning, providing a concrete example and contrasting it with assessment of learning; reviews research demonstrating how enhancing educators' skills in the area of classroom assessment for learning improves student learning; describes what educators need to know and be able to do in order to effectively implement classroom assessment for learning; discusses the most productive way to gain these skills; and outlines the risks of not attending to classroom assessment. (Contains 25 references and 1 table.) (Author)

Assessment for Learning: Classroom Assessment to Improve Student Achievement and Well-Being

By
Judith A. Arter



Chapter 33
Assessment for Learning:
Classroom Assessment to Improve Student Achievement and Well-Being

Judith A. Arter

We educators need to rethink the role of student assessment in effective schools by considering questions such as: What uses of assessment maximize student achievement? How can we best use assessment in the service of student learning and well-being? Do external standardized, high-stakes tests serve us best to maximize the achievement and well-being of the greatest number of students, or is there a better alternative?

We do not normally place the phrases "student assessment" and "student achievement and well-being" in the same sentence. That is because of our own personal experiences with assessment and testing when we were growing up. What feelings do you associate with assessment? Most people associate feelings of anxiety, fear, and nervousness with the idea of assessment, not feelings of eager anticipation, confidence, and well-being. Does assessment have to be like this? Does that kind of assessment environment really maximize learning? Research shows that if we refocus our student assessment efforts away from exclusive concern with large-scale, high-stakes accountability tests, and toward ensuring that every educator has the ability to implement high-quality, student-involved classroom assessment, we can develop far more powerful and nurturing assessment systems. The result will be systems that

- are located where the learning occurs—in the classroom
- are under the control of teachers and students
- empower students to self-assess and self-correct their responses
- leave students looking forward to assessment as a source of information and confirmation, rather than dreading assessment as a source of judgment or control

In other words, we need assessment that not only provides good information to the most important decision makers—teachers and

students—but can also be used to improve the very student achievement being assessed. In this view, assessment will serve us best if we refocus from an almost obsessive emphasis on assessment *of* learning to assessment *for* learning.

Here is how Terry Crooks, a researcher from New Zealand, defines the difference. Assessment *for* learning is roughly the same as *formative assessment*—assessment intended to promote further student learning. Because the intent of this use of assessment is to create more learning, it occurs almost exclusively in the classroom. The phrase "assessment for learning" has become increasingly popular internationally because it better describes this essential use of assessment: in the United Kingdom (e.g., Assessment Reform Group, 1999; Wiliam & Lee, 2001), New Zealand (Crooks, 2001), and the United States (e.g., Shepard, 2000).

Assessment *of* learning, on the other hand, is roughly equivalent to *summative assessment*—assessment intended to summarize student attainment at a particular time. For example, high-stakes, standardized accountability assessments are assessments of learning, as is grading in the classroom. Thus, assessment of learning occurs both in the classroom and through external assessment systems.

We educators are used to thinking about assessment as the measurer of change—as assessment of learning, the index of what students have learned through our various educational innovations. We restructure the school day or put computers in every classroom, for example, then use assessments to see if that made a difference in terms of student achievement. But the concept of classroom assessment for learning presents assessment as the change itself—a direct precipitator of learning, a way to significantly alter the relationships between teachers and students in ways that promote student learning to higher standards. Because improvement in classroom assessment is a change, it is implemented through the change process just as any other change in practice. This has implications for professional development of teachers, leadership, resource allocation, and policy.

Please note that I am not rejecting assessment of learning. It is not that assessment of learning is inappropriate. I just believe it is insufficient to help us reach our goals for student learning. Simply put, we must have a better balance between large-scale and classroom assessment—between assessment of learning and assessment for learning.

This chapter (a) describes in more detail classroom assessment for learning, providing a concrete example and contrasting it with

assessment of learning; (b) reviews research demonstrating how enhancing educators' skills in the area of classroom assessment for learning improves student learning; (c) describes what educators need to know and be able to do in order to effectively implement classroom assessment for learning; (d) discusses the most productive way to gain these skills; and (e) outlines the risks of not attending to classroom assessment.

Assessment for Learning

Assessment for learning—formative assessment—is not a new idea to us educators. During the past several years, however, I have seen new dimensions that take its power to a new level. There has been an explosion of concrete practices and good ideas linked to sound research. (See the sources in the reference list marked with an asterisk.)

Formative assessment is more than testing frequently, although gaining information regularly is important. Formative assessment also involves adjusting teaching to take account of these ongoing assessment results. Yet formative assessment is even more than using information to plan next steps. Here is where the new dimensions come in. Formative assessment is most powerful when students are involved in their own assessment and goal setting.

We involve students in assessment for learning whenever we do things like these:

Help students understand the learning targets they are to reach. What do we want students to know and be able to do at the end of each lesson? Unit? Term? Do students know what we want? After all, which students are more likely to be successful: those who understand the learning targets they are to reach or those who do not? Educators have lots of ways to make learning targets clear to students; examples are using rubrics and scoring guides, stating targets in student-friendly language (e.g., Clarke, 2001, pp. 144–148), and engaging students in determining ways they can tell when they have reached some specified target. (Note: This requires that we, their teachers, also have a clear vision of the learning targets we want our students to reach.)

Engage students in self-assessment. Once students understand the nature of the learning targets they are to reach, they are in a position to begin to evaluate where they are with respect to these targets.

Help students see their own improvement with respect to the learning targets. This happens, for example, with portfolios, where students collect samples of work over time and analyze them for growth.

Give students opportunities to express their understanding. This happens, for example, during dialogue with the teacher or in student-involved parent-teacher conferences where students present the evidence of their own learning (see, e.g., Austin, 1994; Davies & Stiggins, 1996).

Encourage students to set goals and determine the next steps required to move closer to the target.

Such student involvement tends to give students a feeling of control over the conditions of their own success. Research has shown that this control is conducive to learning and results in higher student intrinsic motivation (Caine and Caine, 1997; Jensen, 1998). We all know that one cannot expect positive results from just saying to students, "Now you're going to take control of your own learning through self-assessment and goal setting. So do it." We have to teach students how to do these things.

Royce Sadler (1989, p. 119, as restated in Crooks, 2001, p. 2) discusses what it takes to involve students in their own assessment. First, students must appreciate what high-quality work is. Second, they must have the evaluative skill necessary to compare with some objectivity the quality of what they are producing in relation to the high standard. And, finally, students must have a store of tactics to draw upon to modify their own work. A concrete example of how to accomplish these conditions is the use of rubrics, scoring guides, and performance criteria as instructional tools. Figure 1 outlines strategies for using scoring guides in this manner. Compare the ideas in Figure 1 to the list of requirements described by Sadler. It can be done!

Figure 1. Using Scoring Guides as Instructional Tools
by Judith A. Arter and Jan Chappuis, Assessment Training Institute

What You Need

- A scoring guide (also called a rubric or set of performance criteria) that accurately and completely describes the nature or quality of an important skill, performance, or product you want students to master (e.g., mathematics problem solving, writing, group discussion, oral presentations, science lab reports, literature analysis, critical thinking). A good selection can be found in Arter and McTighe (2001). The scoring guide must be student-friendly and written in language students can understand.
- Anonymous samples of strong and weak student work for the skill or product being taught.

What You Do

1. Teach students the language of quality, the concepts behind strong performance. This step reinforces and validates what students already know, adds to their conceptual understanding of what characteristics contribute to quality work, and ties the terms students use to describe quality to a more formal structure.
   - Ask students to brainstorm characteristics of good-quality work.
   - Show students anonymous samples of low-quality and high-quality work and ask them to expand their list of characteristics based on their examination of these samples.
   - Ask students if they would like to see what teachers think. (They always want to.) Pass out copies of the scoring guide and have them analyze how the features in the scoring guide match with the characteristics they gave.
2. Read (or view), score, and discuss strong and weak sample products or performances. Ask students to use the scoring guide to rate these anonymous samples and justify their rating using wording from the rubric. This process reinforces their ability both to notice what features are important in high-quality work and to use a common vocabulary to describe those features.
3. Use the scoring guide to practice and rehearse making revisions to improve the quality of the work. It is not enough

to ask students merely to judge work and justify their judgments. They also need to understand how to revise work to make it better. Here are various options for doing this:

- Ask students to brainstorm advice for the author on how to improve his or her work. Then ask students (in pairs) to revise the work following their own advice.
- Ask students to write a letter to the creator of the sample, suggesting what she or he could do to make the sample stronger in the dimension of quality under consideration.
- Ask students to rate a product or performance of theirs that they are currently working on, and to revise it to improve the dimension under consideration.

4. Share examples of strong and weak products or performances from life beyond school. Have students analyze these samples for quality using the scoring guide.
5. Model creating the product or performance yourself. Model the messy underside of producing quality work: the initial concepts, how you think through decisions along the way, and what you do when you get stuck. Perhaps ask students to analyze your work for quality and suggest improvements. Revise your work using their advice.
6. Encourage students to share what they know. People consolidate their understanding when they practice describing and articulating criteria for quality. For example, ask students to use the language of the scoring guide to write self-reflections, letters to parents, and papers describing the process they went through to create their work; to revise the scoring guide to make it appropriate for younger students; to write a description of quality as they now understand it (I used to . . . , but now I. . . .); or to participate in conferences with parents or teachers to share their achievement.
7. Design lessons and activities around the dimensions of the scoring guide. Reorganize what you already teach to correspond directly to each dimension of quality in the scoring guide. Make sure students understand how each lesson relates to the scoring guide. This (in addition to step 3) provides students with work-improvement strategies that are keyed directly to the newly learned dimensions of quality.

Note: The "What You Do" section is adapted from work at Northwest Regional Educational Laboratory, Portland, OR.

For additional detail on the strategies outlined in Figure 1 see Arter and Chappuis, 2001 (on applying these strategies to help students develop mathematics problem-solving proficiency), Arter and McTighe, 2001 (applying the strategies to help students become better writers), and Spandel, 2001 (also applying the strategies to writing instruction). Other concrete ideas and practical help on assessment for learning—things you can begin doing tomorrow—can be found in the reference list at the end of this chapter.

Assessment of Learning

The assessment of learning has a long history in this country (Shepard, 2000; Stiggins, 1999b). We began with implementation of the College Entrance Examination Board in the 1930s, and the SAT college admissions test quickly turned into a school accountability measure. Through the 1950s and 1960s we saw the advent of commercially developed norm-referenced, district-level standardized testing programs for local accountability. In the 1970s statewide testing programs made their debut, in the 1980s, national testing programs, and in the 1990s, international testing programs. "Thus we see layer upon layer of tests, each new test expected to accomplish what the prior layers had not done—spark productive school improvement" (Stiggins, 1999b, p. 192). The billions of dollars we have spent on these large-scale, high-stakes assessments of learning is testimony to our national belief that merely by checking achievement status and reporting it we can accomplish important goals:

- Provide the focus to improve student achievement.
- Give all parties the information they need to improve student achievement.
- Apply the pressure needed to motivate educators and students to work harder to improve student achievement.

There is, however, little evidence that these assessments actually acccomplish any of these goals. For example, they do not give teachers and students the information they need to improve student achievement; the results are useful only to those individuals who can use comparable information across students generated once a year—that is, administrators and the general public. Teachers and students make decisions every few minutes, not once a year, so they have to rely on classroom assessments. Annual testing is of minimal value to teachers.

Another mistaken assumption is that these tests motivate educators

and students to work harder so students learn more. Some do, but not all. When faced with what they believe to be unattainable goals or additional public evidence of their failure, some students just give up in hopelessness. Wouldn't it be ironic if the very tests currently being proposed to "leave no child behind"—yearly reading and mathematics tests for all student in grades three through eight—actually were themselves the cause of leaving children behind? Indeed, there is little evidence that large-scale, high-stakes accountability tests have any positive impact on student achievement whatsoever. Robert Linn, a well-known researcher in the area of educational assessment, makes the case strongly: "As someone who has spent his entire career doing research, writing, and thinking about educational testing and assessment issues, I would like to conclude by summarizing a compelling case showing that the major uses of tests for student and school accountability during the past 50 years have improved education and student learning in dramatic ways. Unfortunately, that is not my conclusion" (Linn, 2000, p. 14).

Again, let me emphasize that I am not rejecting assessment of learning entirely, I am only arguing that we need to use it more carefully and to attend more to assessment practices that actually do have a track record of improving student achievement—classroom assessments for learning.

Research on the Impact of Assessment for Learning

Paul Black and Dylan Wiliam (1998) summarized some 250 studies from several countries to answer three questions:

1. Is there evidence that improving formative assessment raises student achievement?
2. Is there evidence of room for improvement?
3. Is there evidence showing how to improve formative assessment?

They reported that "the answer to each of the three questions above is clearly yes" (p. 140). They found that effective use of formative assessment can yield achievement gains of between 0.4 and 0.7 of a standard deviation. This level of improvement translates as follows:

- The typical student in classrooms where formative assessment innovations are taking place would show the same level of achievement as the top 35 percent of students in classrooms where such innovations are not taking place. (A gain from roughly the 50th to the 65th percentile.)

- In a recent international comparison of mathematics achievement, such achievement gains would have raised the standing of a nation in the middle of the pack of 41 countries (where the United States falls) to one of the top five.

The most intriguing of Black and Wiliam's findings, however, was that "improved formative assessment helps low achievers more than other students and so reduces the range of achievement while raising achievement overall" (Black & Wiliam, 1998, p. 141). They further state that the sizes of the effects of improved formative assessment were "larger than most of those found for [other] educational interventions" (p. 141). By way of comparison, a recent article in *Scientific American* (Ehrenberg, Brewer, Gamoran, & Willms, 2001) analyzed the effect of reducing class sizes to fewer than 20 students. In the three best (and largest) studies, reduced class size showed effect sizes that were one half to one third those of improved formative assessment. The authors report that students in smaller classes would gain from 0.05 to 0.2 of a standard deviation. Using the larger number, this would be equivalent to raising achievement from the 50th to at most the 58th percentile.

The upshot of these findings is that if we desire to maximize achievement for all students while decreasing the achievement gap between the highest and lowest achievers, the best solution is to improve formative assessment. "Teachers do not have to choose between teaching well and getting good results" on accountability assessments (Wiliam & Lee, 2001, p. 9). Implementing assessment for learning strategies causes a real improvement in student learning. This improvement is reflected in accountability tests.

Maximizing Formative Assessment

Black and Wiliam also make clear that there is room for improvement in formative assessment as it is commonly practiced. They cite these two specific improvements in formative assessment as being likely to have the biggest impact on student achievement:

- Ensure that classroom assessments yield accurate and important information.
- Give effective feedback.

Let's examine each of these factors in more detail.

Quality of Assessment

Black and Wiliam cite common problems with classroom assessments, such as test questions that emphasize rote and superficial learning and test questions that are poorly written. The heart of the issue, they say, is to make sure that the information generated by classroom assessments is accurate and dependable, so that we can use it to know where students are.

It is not surprising that the accuracy of classroom assessments needs to be improved. Most teachers and administrators have never had the opportunity to learn about assessment. Currently only 14 states require competence in assessment for teacher certification (Stiggins, 1999a), and only 3 states require competence in assessment for principals (Trevisan, 1999). Just think about your own opportunities to learn about assessment. In typical groups of educators only about 5 percent of hands go up in response to the question, How many of you had to take an assessment course to get your certificate? The percentage drops almost to zero when they are asked, How many of you found those courses useful for what you do daily in the classroom? When training programs neglect meaningful assessment competencies, assessment accuracy suffers.

I do not mean to imply that we educators know nothing about assessment. We have had in-service professional development, and we have developed our own expertise through years of experience. On the other hand, in traveling around the country talking to teachers, I have noticed several things. First, although educators are doing some great things with formative assessment, they frequently have trouble articulating why what they are doing is good. In other words, educators tend to have an incomplete understanding of how all the assessment pieces fit with the instructional pieces. Second, educators appear to be doing a lot more assessment of learning, even in the classroom, than assessment for learning. Finally, when they are required to develop assessments, for example assessment systems for accountability (assessment of learning), they become very anxious because of their lack of knowledge.

Although it is not our fault that we do not know as much as we should about assessment, we have a responsibility to learn what we need to know. As Rick Stiggins points out in many of his publications, we are a national faculty untrained in assessment, yet assessment (especially the formative variety) plays an essential role in helping students learn. It is time to do something about this knowledge gap.

Effective Feedback

Feedback is most effective when it is descriptive, is focused on the important learning targets emphasized in the instruction, and includes advice on what the student can do to improve the quality of the work. Maximum benefits occur when students are involved in their own assessment in the ways described previously. Before you start feeling nervous about how much time descriptive feedback will take, I want to relate the results of one research study. Caroline Gipps (2000) compared the effectiveness of different types of feedback to students (nonspecific versus specific) and of the source of the feedback (teachers versus students themselves). Nonspecific feedback (e.g., "you did great," "you need to work on this") made no difference in student achievement. Specific feedback from teachers had a big effect on student achievement (but the teachers were exhausted). Luckily, the largest effect on student achievement came from students giving themselves specific feedback. The moral is that once students have the skills of self-assessment, teachers can save a lot of time by not having to be the sole source of wisdom.

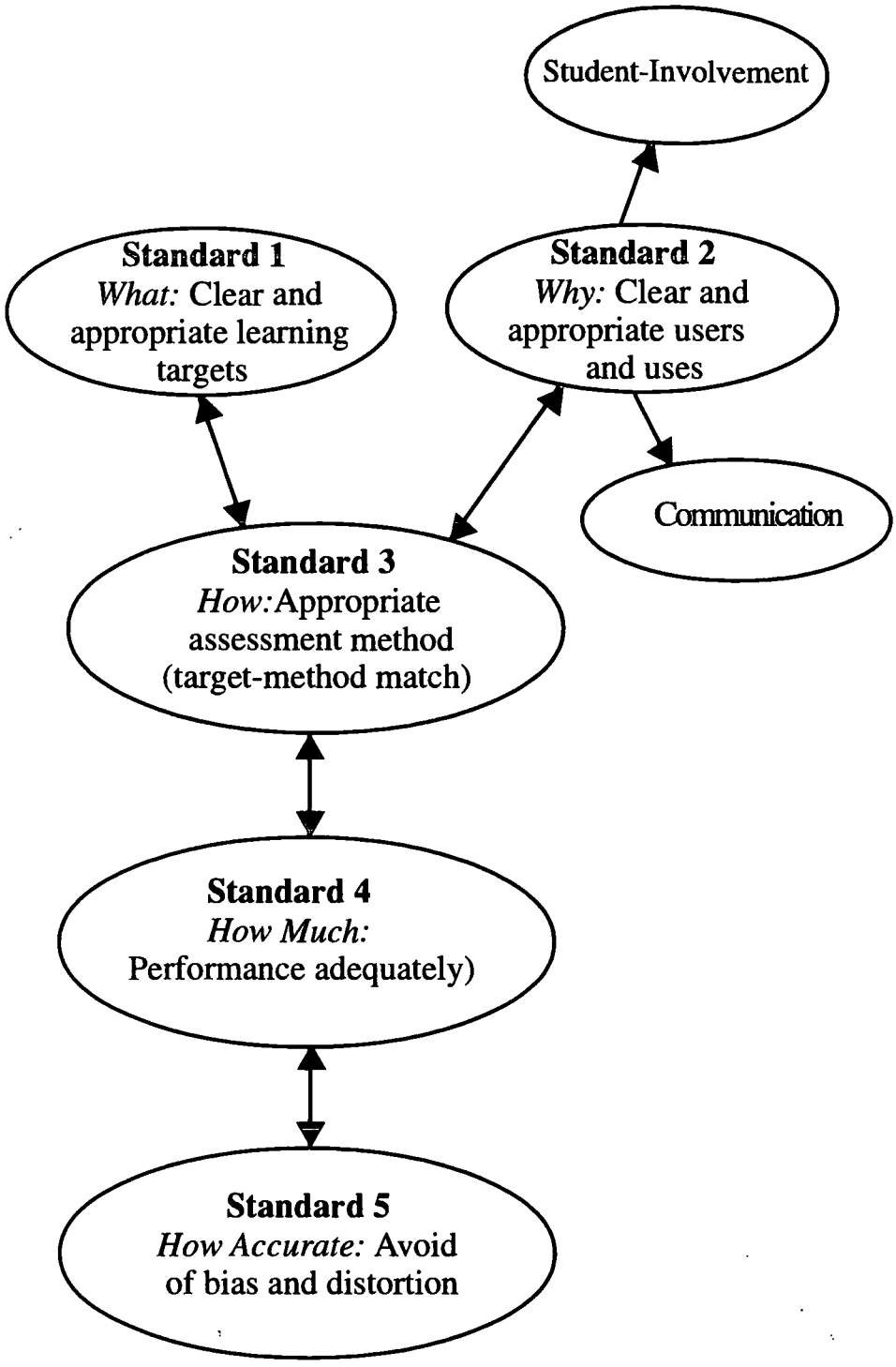
What Educators Need to Know and Be Able to Do

Black and Wiliam's research provides a good outline of what educators need to know and be able to do by way of assessment in order to maximize achievement for the maximum number of students: use accurate assessment, give specific feedback, and involve students. Stiggins (2001) organizes these topics into the structure shown in Figure 2. The structure provides an outline of both the things educators need to know and be able to do, and standards for high-quality classroom assessments. A useful and practical treatment of all five of these standards of quality can be found in the textbook by Stiggins (2001) and the accompanying workbook by Arter and Busick (2001).

Standard 1: What to Assess

First of all, we need to be crystal clear about what achievement targets we want our students to reach. I have already mentioned several times the importance of being crystal clear on learning targets as a prerequisite for formative assessment. Only with achievement targets clearly in mind can we craft instruction and assessment to help students meet those targets. Only with clear achievement targets can we involve students in their own assessment. Moreover these targets have to be appropriate and important.

Figure 2. Five Standards of Quality Assessment



Where do clear and appropriate learning targets come from? They come from research journals; personal content expertise; and several years of effort by states, provinces, and professional organizations that have drafted content standards, or statements of what students need to know and be able to do. With respect to the last, extensive work has been done on tracking standards through grade levels. You may have discovered, however, that these standards and benchmarks usually need additional clarity for use in the classroom. Stiggins (2001, pp. 78–81) provides a very useful way to deconstruct standards into prerequisites that can be more easily integrated into classroom instruction.

Clear targets are also necessary in order to develop accurate, high-quality classroom assessments. Different achievement targets (e.g., mastery of content knowledge, reasoning proficiency, performance skills, and product creation proficiencies) require different types of assessment. The first question to answer is, What do I want to assess? Only with that answer in mind can one determine how best to assess it.

Standard 2: Why Assess

For what purposes are we using assessment? What do we want to accomplish with each classroom assessment? Is the purpose to assign a grade (summative assessment)? Is it to improve learning through student involvement (formative assessment)? Is it to communicate to others the status of student achievement (summative assessment)? Is it to plan the next instructional step (formative assessment)? Is it to provide information to policymakers for accountability (summative assessment)?

These different uses imply different users of the assessment process and its results: parents, students, teachers, and politicians. No single assessment can satisfy the needs of all these people. An accurate and useful assessment is designed with users and uses in mind. For example, parents' need for a summary of their child's learning achievement implies different assessment action than does the need for diagnostic information to plan instruction. Two aspects of users and uses are sufficiently important to be singled out in Figure 2:

Good communication. Different users in different contexts need different information about student achievement in different forms and at different times to do their jobs. Educators need to know who needs classroom assessment information and how to present that information in such a manner that it is clearly understood and can be acted on. A specific example is feedback to students for the goal of improving

learning; for maximum effectiveness such feedback needs to be descriptive, to focus on the learning targets, and to include suggestions on how to improve.

Student involvement. Because students are the most important users of assessment materials and results, they merit a special bubble all their own. We need to understand the relationship between assessment and student motivation. We need to know how to bring students into the process of assessment, thus turning assessment into a powerful instructional intervention.

Standard 3: How to Assess

Whereas the first two standards relate to all three issues of accuracy, feedback, and student involvement, standards 3 through 5 relate primarily to accuracy: how much we can rely on the information garnered from classroom assessments. Specifically, standard 3 relates to understanding assessment methods. Educators need to know how to use the full range of assessment options, including selected response, essay, performance, and personal communication assessment formats. We need to know not only how to write good questions, tasks, and scoring guides, but also when best to use each assessment option. Table 1 provides guidelines for when to use each method.

Standard 4: How Often to Assess

Educators sometimes feel nervous when the concept of sampling arises because it seems highly technical. But sampling is a matter of common sense. It's important to know just how much evidence we need to collect to make a confident conclusion about student achievement. Too much is a waste of time, too little does not provide enough information for good planning. For example, one sample of writing is not enough to determine how well a student writes. One would probably need to sample writing for various audiences and purposes to really know how well students write. Educators do not need to use fancy models of sampling. All they need is a good understanding of the content they are teaching and an awareness that they need to sample all aspects of it.

Table 1. Guidelines for Matching Achievement Targets With Assessment Methods

Achievement Target	Selected Response Short Answer	Essay (Extended Written Response)	Performance Assessment	Personal Communication
Knowledge and Understanding	Strong match for assessing elements of knowledge.	Strong match for tapping understanding of relationships among elements of knowledge.	Depends on the nature of the knowledge to be assessed. Generally not a good match.	Strong match for assessing elements of knowledge and relationships among elements.
Reasoning Proficiency	Can assess some patterns of reasoning in isolation, e.g., main idea, what's most likely to happen next. Other patterns of reasoning, e.g., critical thinking, requires another method.	Strong match. Written descriptions of solutions to complex problem can provide a window into reasoning proficiency.	Strong match. Watching students solve certain problems or examining some products can lead to inferences about reasoning proficiency.	Strong match. One can ask students to "think aloud" or ask follow-up questions to probe reasoning.
Performance Skills	Not a good match. Selected response can be used to assess knowledge about how to do something, but to see if students can actually do it requires a performance assessment.		Strong match	Strong match when the skill to be assessed involves oral communication, e.g., foreign language or oral presentation.
Ability to Create Products	Not a good match. Although selected-response formats can be used to assess knowledge about how to create a product, seeing whether students can actually create it requires a performance assessment.	Strong match if the product involves writing, e.g., a term paper or a poem.	Strong match. One can assess attributes of the product itself.	Not a good match. Can get at knowledge about attributes of quality products but not product quality itself.
Dispositions (Affective Domain)	Strong match. There are lots of published	Strong match.	Possible match. One might be able to infer dispositions from behavior and products.	Strong match. Can talk with students about their attitudes.

Adapted from Stiggins, R. J. (2001). Student-involved classroom assessment (3rd ed.). Upper Saddle River, NJ: Merrill/Prentice Hall, p. 93. Available from the Assessment Training Institute, www.assessmentinst.com.

Standard 5: How Accurately to Assess

This is another topic that tends to put educators on the brink. But, again, the answer is a matter of common sense. Lots of things can be wrong with an assessment, leading to an inaccurate picture of student achievement. We all know this. If any topic was covered in preservice courses on assessment, this was it. Unclear targets and poor matching of targets to assessment methods are two sources of distortion on a test. Other things can be wrong as well: too much writing on a reading test, questions that are not fair to all students, assessment conditions that are not optimal, assessment methods that are not matched to student learning styles, and more. We do not need to understand all the statistical procedures related to test bias, we just need to be aware of the potential sources of bias that can creep into our assessments so that we can do our best to avoid them.

Learning More about Classroom Assessment for Learning

Assessment for learning changes the nature of interactions between teachers and students and constitutes a refinement of instructional practice. If we ask ourselves what features of professional development have helped us the most in changing practice, most educators would

probably cite some subset of the following:

Clear goals. The practices to be learned are clearly framed—we are able to see where we are headed.

Pacing. Learning begins at our current individual levels of understanding and proceeds at a comfortable rate.

Usefulness. The new information or skills we learn quickly deliver benefits in student motivation and learning, saving us time or increasing our confidence.

Practice. We have the opportunity to practice the new ideas or skills in a low-risk environment where it feels safe to stretch.

Collaboration. We have opportunities to work with others to deepen and refine our understanding and application.

Flexibility and efficiency. Learning occurs in a way that fits easily into our diverse and busy schedules.

Long term. We have an extended time to learn and practice.

Not surprisingly, these are the features generally cited in the professional development literature as resulting in real change (Arter, 2001; DuFour & Eaker, 1998). Now consider the professional development options at your disposal: workshops, individual study, and group study. Which options, or combination of options in what proportion, are most likely to provide a beneficial environment for professional development? The evidence suggests that a heavy reliance on individual and group study with limited use of workshops works best. Here is why:

Workshops by themselves can provide small doses of information in an effective and efficient fashion. Experts can sift through the information that participants need to know and can offer motivational sessions that energize an audience so that they want to learn more. But workshops cannot provide the practice with feedback necessary to implement new ideas in the classroom. Individual study by itself allows the learner to tailor information gathering to personal needs, practice with ideas in an applied setting, and proceed at his or her own pace. But individual study can be inefficient and does not allow for practicing

with feedback, bouncing ideas off others, or receiving support during the learning process. Learning teams provide structure for learning about a complex topic, flexibility in structure and pacing, colleagues for group learning, and a support system for practice.

Workshops, individual study, and learning teams are all viable options under certain circumstances. There is considerable agreement, however, that learning teams are the essential element in the mix. Arter and Busick (2001) offer advice on how to set up and conduct professional development learning teams for classroom assessment for learning.

Risks of Not Attending to Classroom Assessment

I have spent considerable ink extolling the potential benefits of refining classroom assessment practices: increases in student achievement and motivation, time savings for teachers, and increases in teacher confidence, to name a few. But I would be remiss if I didn't remind educators of the potential risks of *not* attending to the refinement of classroom assessment.

Classroom Assessment Accuracy

I have already mentioned one potentially huge risk—the risk of inaccurately measuring students. Just think of all the important decisions that teachers, parents, counselors, and students make on the basis of classroom assessment information. For example, students use their classroom assessment performance to make decisions such as these: What should I study? What am I good at? Is it worth trying? Am I capable of learning? Do I like school? Who should my friends be? Is going to college in the picture? What will I be capable of doing in life? Parents make decisions like these: How is my child doing? Is he or she college material? Will I let my child see his or her friends this weekend? Will I support the next bond levy? Teachers make these decisions: What should I teach next? Which students need extra help? Which students should be referred to special programs? How effective is my instruction?Are students working up to their ability? Where are my students performing with respect to state standards?

What would be the impact of making decisions such as these based on inaccurate information? I'm not saying that current classroom assessment information is necessarily inaccurate. I am simply saying that, based on years and years of research on the accuracy of classroom assessments in general, we are currently taking chances that it is. Are

we willing to continue taking that risk, or is it time to refine our practice?

Student Motivation

Another risk involves student motivation. No educator wakes up in the morning and says, "Today I think I'll hurt kids." But the truth is that we have considerable power over student motivation, and in no area do we wield this power more decisively than in assessment. Most of the researchers and practitioners cited in the reference list emphasize the need for basing assessment practices on the recognition that assessment profoundly influences students' motivation and self-esteem, both of which are crucial factors in learning (Black and Wiliam, 1998). Students will not learn if they believe they are unable to learn. Students will not learn if they do not want to learn. Assessment has a large impact on students' decisions about how much they want and are able to learn.

Previous practices in both large-scale and classroom assessment can have inadvertent and unanticipated negative side effects on student motivation. Just remember your reaction to the question, What feelings did you associate with assessment when you were growing up? How many students have gotten the inadvertent message through assessment that they are incapable of achieving, so they might as well give up? How many students build unproductive defensive reactions to school because of the inadvertent but consistent message that they are failures?

According to many authors (e.g., Jensen, 1998; Caine & Caine, 1997), humans have an innate desire to learn; we are born with intrinsic motivation. Learning is required for survival. The brain seeks information, integrates it with other information, interprets it, remembers it, and brings it to bear at the appropriate times. These researchers list the following things as tending to snuff out intrinsic motivation:

- coercion
- intimidation
- rewards and punishments
- negative competitive relationships; comparing one student to another
- infrequent or vague feedback
- limits on personal control
- responsibility without authority

Things that tend to increase intrinsic motivation are these:

- providing a sense of control and choice
- increasing frequency and specificity of feedback
- providing challenge without threat
- encouraging self-assessment

The challenge for us educators is how we can craft the assessment process so it builds student motivation and self-confidence rather than squashes it. Which set of features better describes our traditional methods of large-scale and classroom assessments? Which set better describes assessment for learning? Shepard (2000) documents that assessment has lagged behind instruction in terms of incorporating recent research on how humans learn and what motivates us. It is time to bring assessment into the twenty-first century.

Conclusion

A colleague of mine was teaching a class on portfolios. A special education teacher in this class wanted to try having her students self-assess and track the progress of their writing skills. She was nervous that her students' self-concept would be damaged if they realized how low their performance was, but she tried anyway. She had her students keep selected samples of writing in a folder, learn to assess it accurately using a scoring guide called the "Six Traits + One of Writing" (Spandel, 2001), and describe their progress at the end. She reported that her students scored themselves very low on the scoring guide at first, mostly giving themselves 1s. By the end of the year, their self-assessments were higher, mostly 2s. She also reported that, far from being discouraged, her students were very excited because for the first time in their lives they felt they had control over the conditions of their success—they knew that they had made progress and they knew why.

Assessment for learning can result in assessments that have the following features (Rick Stiggins, personal communication):

- encourage rather than discourage
- build confidence rather than anxiety
- bring hope rather than hopelessness
- offer success rather than frustration
- trigger smiles rather than tears

Stiggins also poses the following question. Answer this question well and you are building the assessment environment stressed in this chapter: What assessments might you conduct next week that your students wouldn't want to miss?

References

The references marked with a star (★) offer practical guidance on implementing assessment for learning

Arter, J. (2001). Learning teams for classroom assessment literacy. *NASSP Bulletin, 85*(621), 53–65.

★Arter, J., & Busick, K. (2001). *Practice with student-involved classroom assessment.* Portland, OR: Assessment Training Institute (www.assessmentinst.com).

★Arter, J., & Chappuis, J. (2001). *Student-involved performance assessment* [Interactive professional development video]. Portland, OR: Assessment Training Institute (www.assessmentinst.com).

★Arter, J., & McTighe, J. (2001). *Scoring rubrics in the classroom.* Thousand Oaks, CA: Corwin Press. Assessment Reform Group. (1999). Assessment for learning: Beyond the black box. Cambridge: University of Cambridge, School of Education.

★Austin, T. (1994). *Changing the view—student led parent conferences.* Portsmouth, NH: Heinemann.

Black, P., & Wiliam, D. (1998, October). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan,* 139–148.

Caine, R. N., & Caine, G. (1997). *Education on the edge of possibility.* Alexandria, VA: Association for Supervision and Curriculum Development.

★Clarke, S. (2001). *Unlocking formative assessment.* London: Hodder & Stoughton.

Crooks, T. (2001). *The validity of formative assessments.* Paper presented at the annual meeting of the British Educational Research Association, Leeds, England.

★Davies, A., Cameron, C., Politano, C., & Gregory, K. (1992). *Together is better.* Winnipeg: Peguis.

★Davies, A., & Stiggins, R. (1996). *Student-involved conferences* [Interactive professional development video]. Portland, OR: Assessment Training Institute (www.assessmentinst.com).

DuFour, R., & Eaker, R. (1998). *Professional learning communities at work.* Bloomington, IN: National Educational Service.

Ehrenberg, R., Brewer, D., Gamoran, A., and Willms, J. D. (2001). Does class size matter? *Scientific American, 285*(5), 78–85.

Gipps, C. (2000). *Classroom assessment and feedback strategies of "expert" elementary teachers.* Paper presented at the annual meeting of the American Educational Research Association, New Orleans.

Jensen, E. (1998). *Teaching with the brain in mind,* Alexandria, VA: Association for Supervision and Curriculum Development.

✦Linn, R. (2000). Assessments and accountability, *Educational Researcher, 29*(2), 4–16.

Sadler, R. (1989). Formative assessment and the design of instructional systems. *Instructional Science, 18:* 99–144.

Shepard, L. (2000). The role of assessment in a learning culture. *Educational Researcher, 29*(7), 4–14.

★Sizoo, B. (2001). *Teaching powerful writing.* New York: Scholastic.

★Spandel, Vicki. (2001). *Creating writers.* New York: Longman.

Stiggins, R. J. (1999a). Evaluating classroom assessment training in teacher education programs. *Educational Measurement: Issues and Practice, 18*(1), 23–27.

Stiggins, R. J. (1999b, November). Assessment, student confidence, and school success. *Phi Delta Kappan,* 191–198.

★Stiggins, R. J. (2001). *Student-involved classroom assessment* (3rd ed.). Upper Saddle River, NJ: Merrill/Prentice Hall.

Trevisan, M. (1999). Administrator certification requirements in student assessment competence. *Applied Measurement in Education, 12*(1): 1–11.

Wiliam, D., & Lee, C. (2001). *Teachers developing assessment for learning: Impact on student achievement.* Paper presented at the annual meeting of the British Educational Research Association, Leeds, England.

✦ Document is included in the Anthology of Assessment Resources CD

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

NOTICE

Reproduction Basis

	This document is covered by a signed "Reproduction Release (Blanket)" form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.
X	This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").