30 Four Principles of Authentic Assessment

Roland Case

There is a common saying in educational circles: "What is counted counts." This expression implies that the truly important learning objectives are those we assess. Student sensitivity to this maxim is implied by their common refrains: "Is this on the test?" and "Will it be for marks?" Consequently, if we value critical thinking or the ability to apply knowledge in new contexts, then we should be concerned that our assessment practices reflect these goals. Unfortunately, most assignments and tests emphasize recall of information. The effect of this is to signal to students that what really matters is remembering facts.

This shortcoming will not be redressed simply by devoting more attention to assessing other goals. Ironically, many ways in which thinking abilities are currently assessed are self-defeating. The "timed" nature of tests and the "once-over and one-time nature" of many assignments do not invite thoughtful student reflection. Advocates of "higher" standards typically call for raised expectations of student performance and for expanded testing. It is not obvious that these steps enhance student learning. High-achieving students who are motivated by grades may already be trying their best, and may be distracted from genuine learning by heightened fears of not doing well on the test. Lesser-motivated students may be doubly discouraged by raising the "educational bar" even further out of their reach and by constantly reminding them of their inferior performance (Assessment Reform Group 2002, 4). In addition, many important educational goals—such as student responsibility, real-life problem solving, reflection, and empathy—are rarely measured. In the rush to "teach to the test," less time may be devoted to these goals. Numerous research studies suggest that many of our system-wide and classroom-based assessment practices inhibit genuine learning.

Overcoming what many regard as the negative effects of common assessment practices is the driving motive for what is referred to as "authentic assessment." The term "authentic" refers to measuring the real, actual, or genuine thing as opposed to measuring a poor substitute. The aim is to supplement traditional assessment practices with "alternative" approaches that offer more meaningful and productive ways of assessing students (Gronlund and Cameron 2004, 10). Although writers describe authentic assessment in varying ways, three interrelated purposes underlie this movement:

- **Greater "authenticity."** Advocates of assessment reform seek a closer fit between the attributes and abilities actually measured by an assessment device and the educational goals that we most value. Too often we assess what is easiest to measure (for example, whether or not students can remember information) and neglect what is more difficult to assess yet nonetheless important (for example, students' ability to think critically and to use their knowledge to solve realistic problems).
- **Supporting learning.** Advocates of assessment reform are committed to using evaluation to help students learn. Often assessment interrupts or discourages learning. We can enhance learning by making assessment tasks more meaningful, by demystifying the process, and by involving students in assessing their efforts and those of fellow students.
- **Fairness to all students.** Advocates of alternative assessment are concerned that some students are penalized by current assessment practices, not because these students know less, but because of the methods and the conditions under which assessment occurs. For example, some students struggle to communicate what they know under the pressure of a single, timed written examination.

In this chapter, I explore four principles for guiding our assessment practices in more authentic ways:

- Focus assessment on what really matters;
- Ensure that assessments are valid indications of student competence;
- Use assessment to support student learning; and
- Develop assessment practices that use the teacher's time efficiently.

In the two follow-up chapters, I discuss ways of nurturing

student ownership of assessment and suggest how to develop and use assessment strategies to further these principles. Before proceeding with these principles, I invite you to assess an assessment device that I used in my own teaching.

Assessing my Assessment

Years ago, after graduating with my teaching certificate, I proudly developed a marking sheet for a research project my students had just completed. Towards the end of a unit on India, I asked students to select an aspect of India (for example, climate, religion, geography, customs) they wished to pursue though independent research. I instructed them to consult several library resources on the basis of which they were to prepare a written report. The report was to include several visuals (for example, charts, graphs, maps) and, unlike previous efforts, these visuals were to clearly connect to ideas in the text. To discourage mere copying of reports from published sources, students were to submit research notes with their final report. When the project was completed, I evaluated and returned their work with a Research Report Assessment sheet (Figure 30.1) attached to the front of each assignment.

> Make a written list of the strengths and weaknesses of my marking sheet. Imagine that I am a student teacher and you are supervising my teaching practicum. Decide the grade you would assign to my assessment practices based on the following scale:
>
> | | |
> |---|---|
> | outstanding | A+/A |
> | very good | A-/B+ |
> | good | B/B- |
> | satisfactory | C+/C |
> | poor | C-/D |
> | very poor | F |

Since developing this assessment device I have asked several hundred pre-service and practising teachers to assess my early effort at assessment. The grades assigned to my marking sheet have ranged the entire spectrum from "outstanding" to "very poor" with the vast majority (approximately 90 per cent of responses) dividing fairly evenly between "good," "satisfactory," and "poor." This variance is cause for some concern. As professionals, how can we have confidence in our assessment practices if there is such latitude in our conclusions about the quality of my marking scheme? This lack of agreement is especially disturbing since our assessments have potentially profound effects on our students. For example, if I was a secondary student and if this assessment was typical of my evaluations, it would have the following consequences:

FIGURE 30.1   RESEARCH REPORT ASSESSMENT

1. Bibliography (1 mark for each book)   /4

2. Notes
    very good (3)
    good (2)
    satisfactory (1)
    poor (0)   /3

3. Charts, maps, drawings, etc.
                    #1   #2   #3
    a) neat:
    b) accurate:
    c) relevant: /9

4. Text
    a) neatness: /2
    b) spelling, grammar, punctuation:
       (1/2 mark off per error)   /5
    c) coverage of major points:
       all (5)
       almost all (4)
       most (3)
       some (2)
       few (1)   /5
    d) well written:
       good (2)
       satisfactory (1)
       poor (0)   /2
                                                          /14

5. Comments:

TOTAL   /30

- "Outstanding" would qualify me for university scholarships.
- "Very good" would enable me to attend the university of my choice, but not on scholarship.
- "Good" would allow me to get into a university, but perhaps not my first choice.
- "Satisfactory" would mean I would be lucky to get into a community college.
- "Poor" would prevent me from directly continuing post-secondary studies.
- "Very poor" would require that I repeat the grade.

Although I do not wish to infer too much from my informal survey, it suggests considerable inconsistency in our understanding of what counts as good assessment. As indicated earlier, I believe there are four principles, which if better

understood and implemented, would improve this predicament. My present purpose is to explain the implications and importance of the four principles. Although there are other principles and other ways of expressing the ones I suggest, the four principles offer a reasonably comprehensive set of considerations for improving our assessment practices.

At the close of the chapter, I will ask you to revisit your initial assessment of my assessment in light of what I hope will be a clearer, more thorough grasp of these key principles. Just as we should use criteria to assess our students' work, so too should we use the principles of authentic assessment as the basis for judging our assessments. The implicit message in my chapter is that we should neither be satisfied with, nor confident in, our assessments of students' work until we have seriously scrutinized our own assessment practices. Let us now look in turn at the four principles that I recommend as a basis for this assessment.

Focus on What Really Matters

The most significant question to ask ourselves when judging our assessment practices is whether or not we are assessing what really matters. Are the criteria we are using—consciously or not—to judge students' work reflective of the most important educational objectives? As suggested above, what teachers assess has important implications for what students consider important and ultimately what they learn. Do our assessment practices do justice to the breadth and complexity of the goals of social studies? Assessments that are skewed towards a limited range of desired outcomes, for example, outcomes related exclusively to factual knowledge, fail to assess and possibly inadvertently discourage student growth along other desired dimensions. This concern is at the root of much of the criticism of standardized testing. Many standardized tests used to evaluate students, teachers, and schools focus on those curriculum outcomes that are easily measured by machine-scoreable questions. This leaves a considerable gap between the outcomes that schools are expected to promote and the outcomes used to measure school performance. In a study from the University of Wisconsin, the overlap between the curriculum and the test for one subject was just 5 per cent (cited in Simmons 2004, 37). The author suggests that the effect is to undervalue some of the most important life skills such as critical thinking and problem solving.

The most shocking realization when I first had occasion to look back at my marking sheet was the imbalance in my assessment. One-third of the total mark for the project (10 marks out of 30) dealt with mechanics (that is, neatness, spelling, and punctuation). Although these are appropriate criteria to use, it now strikes me as mistaken that I would weigh these twice as much as I did the content of the report (the extent to which the report addressed the main ideas accounted for only 5 out of 30 marks). Notice the consequences of this kind of weighting: students who knew a lot about their topic, but who did not write in standard English, might fail the assignment. On the positive side, the fact that I placed some value on information gathering (the use of multiple references and on the taking of competent notes) and on content knowledge (that is, the need for students to cover the main points of the topic) showed some sensitivity to the importance of these goals. Regrettably I did not appear to attach any special importance to the students' ability to think about the material they were researching.

Over the course of a unit or term (not necessarily on any given assignment), we should assess for all relevant goals, and the emphasis assigned these goals should reflect their relative importance. Completing what is called a "table of specifications" is one strategy for checking that each goal is weighted appropriately in one's overall assessment plan. At the end of a reporting period, list all the graded assignments and tests. Record in a table similar to Table 30.1 the amount of marks devoted to each goal. In Table 30.1, the five main goals are indicated in the left-hand column, and the different assessment strategies appear across the top of the chart. The column on the far right reports the percentage of marks assigned to each goal: for example, understanding of key concepts is worth 40 per cent of the total marks (160/400). Be prepared for a surprise when you discover the importance you actually attached to the various goals. The actual weighting of marks should be matched against the importance these goals deserve according to the curriculum and your own professional sense of what really matters, given the students you teach. Although not always possible, setting up a table of specifications beforehand, or partway through a term, allows you to make adjustments for any imbalances in the weighting of certain goals.

Provide Valid Indications

A second consideration in authentic assessment is validity. Although validity has a long history as a complex technical term, in the context of authentic assessment it can be defined as a close fit between the kinds of attributes actually measured by an assessment device and the intended educational goals. In simplest terms, an assessment strategy is valid if it actually assesses the outcomes it claims to assess.

My intention with the marking sheet on the research project was to assess students' ability to identify and use multiple sources of information. I now doubt that assigning a mark for each reference in the bibliography measures this ability. Students could score very well on this part of the assignment even if they did not actually use more than one of

TABLE 30.1 SPECIFICATION OF GOALS ASSESSED

UNIT GOALS	ASSESSMENT STRATEGIES						
	Quizzes	Activity sheets	Group project	In-class observation	Research report	TOTAL MARKS	% OF TOTAL MARK
Critical thinking about issues	15	–	–	20	25	60	15
Information gathering	15	–	20	–	25	60	15
Recall of factual information	50	30	–	–	–	80	20
Understanding of key concepts	20	20	70	–	50	160	40
Co-operation with others	–	–	30	10	–	40	10
TOTAL	100	50	120	30	100	400	100

the books listed in their bibliography. For that matter, I could not be sure that students knew how to find books on their topic—perhaps someone had obtained the books for them. My reliance on the number of references in the bibliography was not a valid indicator of students' research abilities. If I wanted to assess the students' ability to locate and find appropriate sources, I should have created a task in the library where students would be expected to retrieve and assess relevant sources. I could have measured their ability to make use of multiple sources by assigning marks to students who cited several sources of relevant information in their final report. The outcome measured in the "coverage of the main points" section of my marking scheme is equally problematic. Students may have written on all the main points without really understanding what they had put down. If I was serious about finding out if they had gained any understanding of the topic, I would have been better advised to ask students to tell me orally in their own words what they had found out.

The importance of validity was first brought home to me when I was preparing my grade 6 students for a day-long field trip. Several weeks before beginning to plan for a picnic lunch on our field trip, we practised answering word problems like the following:

If there are thirty students in the class and students want on average two sandwiches each, how many slices of bread will be required? How many loaves of bread will we need if there are twenty slices of bread in each loaf? What will be the total cost if bread sells for $1.25 per loaf? How much must each student contribute to cover the cost of the bread?

Despite their ability to successfully solve these kinds of word problems (as determined by a quiz), my students were incapable of determining how much money each would have

to bring for lunch on our field trip. They made no connection between the arithmetic we had been doing and the challenge before them. Even after the connection was explained, they were unable to solve the problem. In the word problems I had provided, all of the mathematical "ingredients" had been supplied to them. Not only did they not know the answers to the real-life questions (that is, the number of sandwiches we would want, the number of slices in a loaf, and the actual cost of bread), beyond getting an adult to tell them, they had no idea how they could come up with the answers. This is one of the dangers when assessment is based largely on isolated assignments and quizzes.

Although I had taught my students to solve word problems on costing lunches, I had not taught them how to cost the lunch. As Grant Wiggins suggests, "school tests make the complex simple by dividing it into isolated and simplistic chores—as if the students need not practise the true test of performance, the test of putting all elements together" (1989, 706). My students' mastery of all the requisite competencies involved in this task and their ability to integrate them successfully were tested only when they were charged with planning the actual lunch. Significantly, I would never have realized the gaps in their abilities, and subsequently addressed them, unless I had assigned this "real-life" assessment task. If we do not assess beyond isolated competencies in artificial situations, we are unlikely to know whether students are able to use their knowledge in significant ways.

Another factor affecting validity is the conditions under which the assessment occurs. The use of "surprise" tests and a failure to make clear to students the basis upon which they will be judged may impair students' abilities to show what they actually know. Instead, students may be rewarded for anticipating what the teacher wants. As well, traditional timed tests reward students who perform well in on-the-spot situa-

tions and may discriminate against students who are equally knowledgeable but are unable to perform under contrived conditions. A very common concern for validity, especially acute with students whose first language is not English, is that students' answers may be a function of their written fluency and not their understanding of the content. Although this obstacle cannot be completely overcome, there may be ways to mitigate its effects:

- Assignments and questions should be explained orally to students, and perhaps have someone translate the instructions and make frequent use of visual aids and other low vocabulary prompts.
- Whenever feasible, allow students to represent their answers in graphic form, orally, in written point form, or perhaps even in their native tongue.
- Whenever appropriate, offer alternative assignments, reduced expectations, or additional assistance to offset any language impediment.

Besides being careful when developing measures to devise questions or tasks that capture what we intended, validity may also be enhanced by using several devices of different sorts to gather information about student achievement. The point of considering a variety of approaches is to increase the likelihood of finding a valid way to assess the desired outcomes. If, for example, the ability to solve real-life problems is an important goal then, at some point, we should assess the students' ability to act on a real problem and not be satisfied by asking students to list the factors they would consider in a hypothetical context. An observation checklist or rating scale may be particularly effective in assessing student performance in group projects and class presentations. Having students keep a journal while participating in a project or a simulation activity may provide rich information about student attitudes towards themselves and others. For example, while preparing for a class discussion or debate, students might comment on their reflections about expressing and defending their positions, or about working with others.

Use Assessment to Support Learning

Advocates of assessment reform are emphatic about using assessment to enhance learning. The enhanced emphasis on using assessment to support learning is reflected in the distinction between the traditional phrase "assessment of learning" and the more recently introduced notion of "assessment for learning" (Assessment Reform Group 1999, 2). In their review of numerous studies, this group concludes that students would be better motivated and learn more if assessment

practices focussed more on supporting learning than on measuring learning (Assessment Reform Group 2002, 10). More recently, educators are talking about assessment as learning to heighten awareness of the potential to use assessment tasks as opportunities for learning, not simply to provide formative feedback (British Columbia Ministry of Education 2005, 23-24). Self-assessment is an example of an assessment task that is also a learning task as students examine their own work and think through its strengths and shortcomings.

Greater validity of assessment measures is in itself an attempt to use assessment to support learning. As suggested by the example about planning for the field trip lunch, if an assessment does not capture what it is we really value, then we are less likely to know when we have succeeded (or have failed to succeed) in reaching our objective. Only after the real-life task did I realize that my students could not calculate the cost of our lunch. Assessment practices can support learning in at least four other important ways:

- clearly communicate expectations;
- involve students in the assessment process;
- provide helpful feedback on learning; and
- provide opportunities and incentives for students to improve.

COMMUNICATE EXPECTATIONS

If students know clearly what is expected of them they are more likely to succeed at the task. One of the most obvious ways in which I could have used my assessment practices to support learning was by presenting students with the marking sheet before they embarked on the research assignment. As it was, they saw the criteria only after they had completed their report. If my measure had had validity and had focussed on the important goals, I would have been signalling to students what was important and what they were required to do to demonstrate their learning. But because of its flaws, had I distributed my original marking sheet beforehand, unwittingly I would have been encouraging students to attend to the technical dimensions more than the content. The fact that I instructed students to select graphs, charts, and maps that related to their text and that I assessed for this, encouraged students to attend to this feature in their reports.

Students may be even clearer about expectations if they are informed specifically about the "criteria" upon which they will be marked and the importance of those criteria (that is, the number of marks assigned to each criterion) and the "standards" for achievement of these criteria. Because the concepts of criteria and standards are often used interchangeably, let me explain the distinction drawn between these two terms.

Criteria are the features or attributes that provide the grounds for judging quality. Sample criteria include:

- historical accuracy
- originality of ideas
- use of several sources
- clarity of presentation
- depth of answer
- active participation in project
- openness to new ideas
- flow/structure of the paper
- neatness
- spelling accuracy

Standards are the benchmarks, performance levels, or degrees of achievement of a given criterion (that is, "high" and "low" standards). Standards can be binary (for example, correct/incorrect, pass/fail, satisfactory/unsatisfactory) or have multiple levels (for example, A+ to F, outstanding to very weak, well above expectations to not yet meeting expectations). Sample standards for three criteria are listed in Table 30.2.

My grade 6 students might have been better able to succeed had I clearly indicated all the criteria and standards for assessment. When assessing their notes I merely indicated whether they were "very good," "good," and so on, without indicating the basis for this assessment. What criteria was I using? Was it the neatness of the notes? Conciseness? Amount of notes? Or, perhaps all of these? Furthermore, even if students knew the criteria, they may still not know what distinguished a "good" from a "satisfactory" standard of note-taking. And yet, if I wanted them to improve, this is precisely the understanding they require. I did a slightly better job of communicating the criteria and standards for the "main points": my criterion was the amount of coverage and my standards were distinguished by the number of main points covered (for example, all, most, a few). Besides supporting learning, another powerful reason for clearly articulating standards is that it reduces inconsistency and arbitrariness in assessments. I now wonder when I look at the standards I offered for "coverage of main points" if there is any real difference between "almost all" and "most" points and between "some" and a "few" points. If there is no clear distinction between these performance levels, how can I have reliably distinguished among them?

In the spirit of living the principles I preach, I offer in Figure 30.2 detailed descriptions of performance levels or standards for each of the four principles that I offer as the criteria for judging authentic assessments. After you have finished reading about all four of these principles, and have a clear understanding of what each involves, I will ask you to use this assessment rubric to reassess your original judgment of my marking sheet. For the time being, I offer this as an example of a way in which we can support learning by clearly articulating the standards for our assessment criteria. Read my descriptions of each standard and decide if you would recognize what each involves.

TABLE 30.2 SAMPLE STANDARDS		
CRITERION	**STANDARDS**	**DESCRIPTIONS OF PERFORMANCE LEVELS**
Historical accuracy	excellent	no factual inaccuracies
	good	at most, a few minor factual inaccuracies that do not affect the conclusion
	satisfactory	one major inaccuracy and several minor factual inaccuracies
	unsatisfactory	several or more major factual inaccuracies that completely undermine the conclusion
Depth of answer	in-depth	all main topics are analyzed in a probing and careful manner
	modest depth	although there is evidence of careful analysis, some aspects are not addressed in much depth
	superficial	for the most part, topics are not addressed superficially
Spelling accuracy	excellent	zero errors
	very good	at most 2 errors
	good	between 3 and 5 errors
	satisfactory	between 6 and 9 errors
	poor	10 or more errors

FIGURE 30.2   ASSESSING THE ASSESSMENT

	HIGHLY EVIDENT	MOSTLY EVIDENT	PARTIALLY EVIDENT	COMPLETELY ABSENT
Focusses on the important goals	6 The weighting of marks closely matches the important objectives of the assignment.	4 The weighting of marks generally matches with the important objectives of the assignment.	2 The weighting of marks is out of balance with important objectives of the assignment.	0 The weighting of marks misses or seriously under-represents all the important objectives of the assignment.
Provides valid indications of student ability	6 The assignment and the marking scheme directly measure student ability on all intended outcomes.	4 The assignment and the marking scheme measure in a fairly direct way student ability on important intended outcomes.	2 The assignment and the marking scheme are unlikely to measure student ability on some of the key intended outcomes.	0 The assignment and the marking scheme measure student ability in a superficial, contrived, or distorted manner.
Supports student learning	6 The device very clearly identifies the criteria and standards and provides very helpful feedback for improvement. Has significant potential to reinforce and encourage important student learning.	4 The device is generally clear about the criteria and standards and provides some helpful feedback for improvement. Has some potential to reinforce and encourage student learning in some major areas.	2 The device contains significant gaps or ambiguities in communicating the criteria and standards and offers little helpful feedback for improvement. Key aspects of the assessment fail to reinforce and encourage student learning.	0 The device is very vague or confused about the criteria and standards, and offers no helpful feedback for improvement. Offers nothing to support, and may discourage, significant learning.
Uses teacher time efficiently	3 The assessment and feedback method very efficiently uses teacher time in providing significant information to students.	2 The assessment and feedback method is somewhat efficient in its demands on teacher time relative to the rewards.	1 The assessment and feedback method is somewhat inefficient in its demands on teacher time.	0 The assessment and feedback method requires very extensive teacher time relative to what it communicates.

Outstanding (A+/A):   19–21
Very Good (A-/B+):   16–18
Good (B/B-):   12–15
Satisfactory (C+/C):   9–11
Poor (C-/D):   5–8

Total: /21

Grade:

INVOLVE STUDENTS IN ASSESSMENT

Involving students directly in the assessment process is another way to support learning. The next chapter in this collection, "Building Student Ownership of Assessment," has more to say on each of the following areas for student involvement:

- **Setting criteria and standards.** Joint teacher and student negotiation of the criteria upon which students are to be judged increases student understanding of what is expected and ultimately of their performance in light of these expectations. Students can also be involved in deciding upon standards—by articulating what might be required in order for the work to be regarded as excellent, good, and so on.
- **Creating assessment tasks.** Another way to involve students is by inviting them to assist in developing the tasks upon which they will be assessed.
- **Self- and peer assessment.** Involving students in self- and peer assessment can greatly enhance their learning. The very exercise of assessing their peers on the specific criteria related to the lesson would likely reinforce the students' own understandings of what is expected of them. Furthermore, involving students in assessment encourages students to take greater ownership of their learning. An important dimension of self-assessment is communicating the results to others—either to the teacher, their peers, or parents.

PROVIDE FEEDBACK ON LEARNING

We can enhance learning by helping students see how they might improve. Providing students with useful feedback must go beyond assigning a mark or offering a brief summative comment. For example, in my marking scheme, I provided students, albeit after the fact, with a detailed breakdown of how well they did on each aspect, with a place for general comments. This helped them understand what they did well and where more attention was needed.

Although some students may be concerned exclusively with their mark on an assignment, Paul Black and Dylan Wiliam (1998) have found that this leads to no improvement in student achievement—marks are entirely about assessment of learning, not assessment for learning. In fact, they conclude that grading and other forms of comparative feedback actually get in the way of learning and are especially demotivating for less accomplished students. Because of the negative effects of repeated failure, some educators recommend providing scores only at the end of the grading period when it is necessary to prepare an evaluation report. According to the Assessment Reform Group (2002, 10), ideally teachers should assign marks only if students have a good chance of succeeding. In the interim, students should be provided with abundant feedback and encouragement.

If we want students to improve, our feedback must clearly communicate what has been successfully done, where improvement is needed, and how to do this. A carefully prepared rubric can go a long way in providing this feedback, both in terms of indicating how students have done and what might be done to improve their performance. In my own experience, students benefit most from the use of rubrics when marks are not indicated. The lack of a summative judgment requires them to read the descriptors more carefully and encourages them to believe that it is not too late to improve. Other methods of providing effective feedback include:

- very specific written teacher comments;
- teacher conferences;
- comments by fellow students explaining areas for improvement;
- large and small group discussion of answers; and
- exemplars—samples of high-quality performance—of student work, so long as improvement requires more than simply copying the ideas in the exemplar.

PROVIDE OPPORTUNITIES AND INCENTIVES TO IMPROVE

Where feasible, use assessment to encourage students to learn on their own and to revise and rethink their work. Possible strategies include establishing a habit of assessing key objectives in subsequent units, and making it clear to students that certain abilities will be assessed routinely. Some students may be motivated by supplemental tests or makeup assignments for those who make some effort to improve their understanding. One of my most counter-productive assessment habits as a public school teacher was my penchant for "one-shot" efforts. Rarely did I ask students to seriously revise their work—if work was revisited it was only to tidy up typos or add a missing sentence or two. Now, in my university teaching, I no longer have one-time assignments. In my graduate class, for example, instead of writing three different papers, my students write the same paper three times. The first and second drafts are distributed to everyone in the class for critique. In the first draft, students show largely what they could do before the course. The significant improvement—the deeper, more insightful learning—occurs with the two subsequent revisions where students work through the ideas raised by their colleagues and by me.

Before inviting students to undertake serious revision, we should ensure that they have meaningful input as to how they did initially and what they might do to improve. Since

elementary and secondary students may be less motivated to engage in subsequent revisions than are students in graduate school, we must encourage them in this regard:

- Ask students to redo only a part of the original assignment (for example, the two worst [or preferred] answers, or the opening and closing paragraphs of an essay).
- Create additional incentives for revising a draft (for example, revised assignments might be exhibited in a fair, submitted to the newspaper, published in a book, or otherwise shared with adults or other students).
- Comment on but do not mark the initial mandatory draft. Establish that only the revised draft "counts" for marks.
- Ask students weeks or months later to revisit an earlier work to see how much they have progressed in the intervening time.

When encouraging students to learn from feedback, it is not simply a matter of them redoing completed assignments, but also formulating plans to use the lessons learned to improve upcoming projects. For example, we might ask students to identify a learning goal, anticipate an obstacle they might face, and suggest how they might overcome it.

Using Teachers' Time Efficiently

The final, perhaps one might say the bottom-line, criterion of good assessment is efficient use of teacher time. Although efficiency has no direct relationship to authentic assessment, the incredible press on teachers' time means that changes, however desirable, are unlikely to occur if they are more time-consuming. Generally speaking, marking sheets, including the one I developed for the independent research project, are efficient assessment tools. Once familiar with the layout it is easy to complete the sheet quickly because it keeps the assessor focussed and saves having to repeatedly write out the same comments. Rubrics are great savers of marking time, but they require considerable up-front development time. For this reason, I am inclined to develop rubrics for major projects during the year—starting with the one that causes the biggest marking headache—and when I want students to undertake peer or self-assessment.

Clearly articulated criteria and standards, communicated beforehand, increase the likelihood of students providing what the teacher is looking for, and help focus the teacher's attention when marking assignments. Clear expectations reduce the likelihood of protracted discussions with students who complain that they did not know what was required of them.

Student peer and self-assessment can save teacher time provided students are adequately trained in the practice. It saves time because it means that students are giving each other feedback that otherwise the teacher would have to give. Developing students' abilities to assess their own work and their peers' work may be one of the more efficient "learning" strategies. In my university teaching, I marvel at how much graduate students learn about (and improve upon) their own writing from frequent opportunities to critique the work of fellow students. They are better able to appraise their own writing after noting the similar strengths and weaknesses in others' writing and they benefit considerably from other students' critiques of their own work. But perhaps the biggest efficiency arising from self-assessment comes from a shift in the perceived ownership of learning. When students truly realize that they, and not the teacher, have primary responsibility for the grade they receive, the relationship between student and teacher changes. There is less need for the teacher to chase after the students and drum the information into them. Students acquire more independence, self-reliance, and commitment—and, and as a result, more is learned.

A Final Reflection

Return to your initial assessment of my marking sheet. While reviewing any notes you took, consider the merits and oversights in your earlier thoughts about my device. Use the assessment rubric presented earlier to reassess my marking sheet. What grade do you now think that it is worth? Even if your assessment is largely unchanged, do you now have greater confidence in the grade you assigned? Is it a fairer, more valid assessment? Are you clearer about how you might help me improve my assessment practices? I hope the answer is yes to all these questions, and to one further question: Do you have a better understanding of principles to follow in making your assessment practices more authentic?

Select an assessment device (for example, a quiz, end-of-unit project, an observation checklist) that you have developed or that is included in a teaching resource. Use the rubric "Assessing the Assessment" to evaluate this device. Based on what you have learned about the four principles discussed in this chapter, suggest ways to make the device more authentic.

ACKNOWLEDGMENT

REFERENCES

Assessment Reform Group. 1999. *Assessment for learning: Beyond the black box.* Cambridge, UK: University of Cambridge. Available at http://arg.educ.cam.ac.uk/publications.html.

———. 2002. *Testing, motivation and learning.* Cambridge, UK: University of Cambridge. Available at http://arg.educ.cam.ac.uk/publications.html.

Black P. and D. William. 1998. Inside the Black Box: Raising Standards Through Classroom Assessment. *Phi Delta Kappan* 80 (2): 139–148. Available at http://www.pdkintl.org/kappan/kbla9810.htm.

British Columbia Ministry of Education. 2005. *Social Studies 10: Integrated Resource Package 2005* (Response Draft). Victoria, BC: Author.

Gronlund, N.E. and I.J. Cameron. 2004. *Assessment of student achievement* (Canadian edition). Toronto: Pearson Education Canada.

Simmons, N.E. 2004. (De)grading the standardized test. *Education Canada* 44 (3): 37–39.

Wiggins, G. 1989. A true test: Toward more authentic and equitable assessment. *Phi Delta Kappan* 70: 703–713.