1

Basing Grades on Standards

In a truly coherent system, grading and reporting practices must align with the principles of standards-based reform.

—Shannon & Bylsma, Office of Superintendent of Public Instruction (Washington), p. 79

GUIDELINE 1

Relate grading procedures to learning goals.

a. Use learning goals (standards or some clustering of standards [e.g., strands]) as basis for grade determination and grade reporting.

b. Use assessment methods as the subset, *not* the set (i.e., standards, Learning Results, Expectations, Outcomes, etc.).

The Case of . . .

Elliot's Amazing Passing Grade

In Grade 9, in a program that introduced students to the wide range of possibilities open to them, Elliot was required to take a course in a vocational area. Elliot chose auto mechanics, even though he had very little interest or skill in this area. During

(Continued)

(Continued)

the six weeks of the class, he completed two poor-quality repairs of simple problems, both of which deservedly received very low marks. School procedures established a highly structured assessment schedule, which provided four days of written exams at the middle and at the end of each semester. School policy also required that a single subject grade be reported for each subject and that exams be held in each subject with the score counting for 50 percent of the final grade. The auto mechanics exam included questions about safety procedures and how to make simple repairs. This assessment was easy for Elliot because he had a good memory and wrote well. Elliot received 50 out of 50 on the exam, which was added to his performance marks. This combination resulted in an overall passing grade, which Elliot clearly did not deserve, as the main goal of the course was for students to perform quality repairs.

WHAT'S THE PURPOSE OF THE GUIDELINE?

This guideline requires that grading procedures be aligned with stated learning goals. This alignment is direct, and ideally a grade is determined and reported for each learning goal with no overall grade. As is illustrated in the case study, Elliot should have received a low grade for practical skills and a high grade for knowledge; combining his achievement on these two different learning goals into a single grade is both difficult and meaningless. As Tombari and Borich (1999) note,

> The principal limitation of any grading system that requires the teacher to assign one number or letter to represent . . . learning is that one symbol can convey *only* one meaning. One symbol *cannot* do justice to the different degrees of learning a student acquires across all learning outcomes. [emphasis added] (p. 213)

However, in many schools, especially middle and high schools, for the foreseeable future, teachers will be required to determine single-subject grades. Where this is the case, the contribution of each learning goal to the final grade needs to be clear and direct. For example, if the primary learning goal in a course is practical demonstration of skills, then the final grade in that course should be based mostly on direct observation of those skills and evaluation of the products that result from those skills. Teachers record keeping, therefore, must be based on learning goals, not assessment methods.

This guideline requires that grading procedures be aligned with stated learning goals. This alignment is direct, and ideally a grade is determined and reported for each learning goal with *no* overall grade.

WHAT ARE THE KEY ELEMENTS OF THE GUIDELINE?

As noted earlier, most schools and school districts and all states and provinces now have clearly stated learning goals. Different words are used to describe these

goals. In most places, *standards* is still the descriptor of choice, but in many places, other words, such as *learning results* or *expectations,* are used. It does not matter much which word is used; the concept is that at either the local or state level, specific learning targets have been established, often on a grade-by-grade basis. In this chapter, to simplify a confusing situation, I use *learning goals* as a generic term; however, when other sources are quoted, alternative terms to *learning goals* will be retained.

Learning Goals

Grades should be effective communication vehicles, and the methods used to determine them need to provide optimum opportunities for student success and to encourage learning. For this to happen, the meaning of grades must be clear, which requires that, in addition to all the issues dealt with in the other guidelines, grades must be directly related to the learning goals for each grading period in each classroom. Teachers must understand clearly what learning results are expected and then base their assessment and grading plans on these learning goals. Students must also understand clearly what the learning goals are so that they know what is expected of them.

The need for a learning goals base is clearly illustrated by Winger (2005):

> In the middle of the semester, [a colleague] asked her language arts students to identify one area in which they hoped to improve during the second half of the course. Instead of identifying a skill, such as writing organization or reading comprehension, most students listed tests or homework. Rather than identifying gaps in student learning, this teacher's grading practices had focused students' attention on the assessment tools. (p. 62)

Winger goes on to state that "a grade that is separated into distinct components on the basis of key learning becomes a meaningful communication . . . about what students have mastered and not mastered" (pp. 62–63). He believes that he developed "a healthier grading system" when he organized the learning goals in the Introduction to Sociology class for juniors and seniors into four categories—conceptual understanding, application, analysis and evaluation, and formal writing—and based his grading and reporting on these four categories. Likewise, Susan Christopher (2007/2008), a middle school Spanish teacher in Clayton, Missouri, uses "the major skills that I want students to know and do: understand written and spoken Spanish, write and speak comprehensibly, and accurately use the vocabulary and grammar structures that we've learned."

Grading Plan

Off Target: Methods of Assessment

Before discussing an appropriate basis for determining grades, let's briefly discuss what not to use. Simply put, do not base a grading plan on methods of assessment, as illustrated in Figure 1.1.

Figure 1.1 Traditional Grading Plan

Evaluation Category	Expected Range
1. **Quizzes/tests/exams**	20–30%
2. **Written assignments** creative or explanatory paragraphs, essays, notes, organizers, writing folios, portfolios	15–25%
3. **Oral presentations or demonstrations** brief or more formal presentations or demonstrations, role-playing, debates, skits, etc.	15–25%
4. **Projects/assignments** research tasks, hands-on projects, video- or audiotaped productions, analysis of issues, etc.	10–20%
5. **Cooperative group learning** evaluation of the process and skills learned as an individual and as a group member	5–15%
6. **Independent learning** individual organizational skills, contributions to class activities and discussions, homework, notebooks	5–15%
	70–130%

NOTE: Aspects of this plan conflict with other grading guidelines in addition to Guideline 1.

Grades must be directly related to the learning goals for each grading period.

With this type of plan, it is extremely difficult to emphasize each learning goal appropriately because the primary focus is on the methods of assessment. Each learning goal may be assessed in a number of ways; for example, there may be questions on tests/exams, written assignments, and demonstrations for each goal. However, to align assessment with the desired emphasis on each goal over several methods of assessment is extremely difficult.

On Target: Learning Goals

A much better approach is to use the learning goals as the basis for grades. In this approach, some aspect of the organizational structure of the learning goals is the basis for grades for the year or for each grading period. This can be determined by teachers working collaboratively; for example, all the Grade 3 teachers in a school or district or all the Grade 9 science teachers meet to discuss what is the most appropriate basis for grades. This discussion may be the best professional dialogue teachers engage in, because they have to be very clear about what goals are important at what point in the school year and they have to be prepared to support their own views while respecting the opinions of others. Another important benefit of this approach is that much greater consistency across a school or district will occur than with traditional, largely private approaches to grading.

Ideally, the organizational structure chosen should be at the most specific and detailed level of the learning goals because, as Marzano (2000) points out,

"A problem that makes the traditional system highly ineffective . . . is the mixing of different types of knowledge and skills" (p. 13). He further points out that "the construct of academic achievement is not a simple one" and "to provide effective feedback to students, teachers must keep track of those factors they wish to include in grades" (p. 40). Thus, the most appropriate way to organize a grading plan would be to base it on individual standards or benchmarks.

This is being done with increasing frequency by elementary schools and is reflected in a report card such as the ones in Figures 11.2a and 11.2b. To complete such a report card effectively, each teacher would need to use a tracking sheet for each student with horizontal rows for each of the standards included on the report card (see Figures 1.2–1.6). Unfortunately, in many jurisdictions, there are too many standards and/or teachers have too many students to manage tracking of every standard for every student, so they must find a compromise. This often means using the strands in the content standards or the basic organizing structure from the performance standards. Examples of such approaches and different levels of specificity are provided in Figures 1.2 through 1.6. Until recently, the state of Washington had five strands (and many standards) in its state goals for mathematics. The strands were as follows:

1. Concepts and Procedures
2. Problem Solving
3. Reasoning
4. Communication
5. Connections

Figure 1.2 shows these strands and a tracking sheet that a teacher could use to record the achievement evidence for each strand. Each component of every assessment had to link to one of the five strands. A grade was then determined for each strand, and if necessary, an overall grade could be determined for mathematics. The former is more desirable as it provides more useful information, but the approach taken would depend largely on the structure of the report card and the policies/procedures of the school or district.

The "Concepts and Procedures" strand included standards in Number Sense, Measurement, Geometry, Probability and Statistics, and Algebra. It is highly likely that teachers, students, and parents would it find it helpful to have information about student achievement in each of these groupings. Figure 1.3 on page 51 shows an approach and a tracking sheet that would be appropriate for this greater specificity. In addition, the way assessment results were recorded is shown. For example, on September 1, a test was administered that included standards from Number Sense, Algebra, Reasoning, and Connections, so scores were recorded for each of those categories. Then, on September 8, a performance assessment was completed that included standards from Measurement, Algebra, Problem Solving, and Communication, with rubric scores (on a four-point scale) recorded for each category. Over the grading period, this approach would be continued, so at the end, grades for each category assessed during the grading period could be determined and, if necessary, an overall report card grade could be determined. (See also Figure 1.4 on page 52 for an example of a tracking sheet that has been filled in for an entire grading period.)

Figure 1.2 Washington State Math Strands

Summary of Evidence for WA Mathematics EALRs*

Student:_____

	Achievement Evidence												
Assessments ▶ **▼Standards**													Summary
Concepts and Procedures													
Problem Solving													
Reasoning													
Communication													
Connections													
Comments:								Most consistent level of achievement with consideration for more recent					

* Essential Academic Learning Requirements

Figure 1.3 Washington State Math Strands With Concepts and Procedures

Summary of Evidence for WA Mathematics EALRs*

Student:_____

	Achievement Evidence												
Assessments ▸ / **Standards/ ▾Strands**	9/1 TEST	9/8 PA											Summary
Concepts and Procedures													
1. Number Sense	**9/10**												
2. Measurement		**3**											
3. Geometry													
4. Prob. and Stats													
5. Algebra	**9/10**	**3**											
Problem Solving		**4**											
Reasoning	**9/10**												
Communication		**4**											
Connections	**9/10**												
Comments:													
								Most consistent level of achievement with consideration for more recent					

* Essential Academic Learning Requirements

Figure 1.4 Washington State Math Strands With Concepts and Procedures (Complete for Grading Period)

Summary of Evidence for First-Quarter WA Mathematics EALRs*

Student:_____

	Achievement Evidence												
Assessments ▶ / **Standards/ ▼Strands**	9/1 Test	9/8 PA	9/15 Test	9/22 PA	9/29 PA	10/6 Test	10/17 PA	10/24 PA	10/27 Test			11/1 Exam	Summary
Concepts and Procedures													
1. Number Sense	**9/10**		**9/10**	**4**			**4**					**19/20**	**A**
2. Measurement		**3**	**8/10**		**3**		**3**					**16/20**	**B**
3. Geometry								**4**					**I**
4. Prob. and Stats	**6/10**	**1**		**1**		**6/10**						**13/20**	**D**
5. Algebra									**10/10**				**I**
Problem Solving		**4**	**4**				**4**					**4**	**A**
Reasoning	**9/10**			**4**				**4**				**4**	**A**
Communication		**4**			**4**		**4**					**4**	**A**
Connections	**3/10**				**1**	**7/10**		**4**	**19/20**			**4**	**A**
Comments:	**A = 90–100% = 4** **B = 80–89% = 3** **C = 70–79% = 2** **D = 60–69% = 1** **F = 0–59% = 0** **I = Incomplete/ Insufficient Evidence**												
									Most consistent level of achievement with consideration for more recent				

* Essential Academic Learning Requirements

It is almost self-evident that standards-based grading also requires standards-based assessment; if tests are used, each test must be on a single learning goal or the test must be organized by learning goals (or concepts or skills). For example, on a pre-algebra test developed by Forrest Clark and his colleagues at Nisqually Middle School in the North Thurston (Washington) School District, questions 1–3 were identified as being on adding and subtracting integers, questions 4–8 on multiplying and dividing integers, questions 9–18 on solving two-step equations, questions 19–23 on plotting points, and questions 24–27 on powers of 10. Based on their answers, students were then identified as "mastered," "improving," or "needing help" on each concept.

This type of approach has also been adopted by a California mathematics teacher, Dan Meyer. On his blog (2007), he says that for him,

> learning breaks across *skills* not chapter units. Instead of assessing at the end of chapters, we assess at the completion of a significant skill. Instead of lumping all the skills together under one grade (making that grade useless beyond a "did good" or "did bad" level) we track each skill separately in our grade book.

As a result of this approach, Meyer says,

> Students *like* assessment. . . . Students *like* the process. They *know* which skills they need to improve (because we track them separately—me and them, both), they *know* how they can improve them (by studying or coming in for tutoring), and they *know* they ll be rewarded for their efforts.

It is important to emphasize that in the Nisqually example, no overall score for the test was recorded or reported. The traditional practice of only recoding an overall score provides very little information of value and should be discontinued. On a test of 40 questions like this one, a number of students might get 30 out of 40 correct with each having a very different pattern of performance on the four learning goals. The valuable information is the profile—the details of the achievement on each learning goal—not the overall score. As Melograno (2007) states, "It should be clear that single-letter grading is incompatible with the meaning of standards-based education" (pp. 47–48).

There were also subcategories for the other four Washington mathematics standards, so some teachers/schools might have wanted to grade at the level of specificity indicated in Figure 1.5. At this level of specificity, there would have been 5 strands and 17 subcategories.

The issue of the level of specificity at which to collect and report evidence of student achievement is very complex, as we have to try to balance the amount of information with workload for teachers and possible information overload for parents. Marzano (2006) provides a useful approach to solving this dilemma in his book *Classroom Assessment and Grading That Work.* He suggests that teachers—who are the subject and grade-level experts—should select "no more than 20 measurement topics [or learning goals] per subject, per grade level and ideally about 15" (p. 23).

Figure 1.5 Washington State Math Strands Expanded

Summary of Evidence for WA Mathematics EALRs*

Student:_____

	Achievement Evidence												
Assessments ▶													Summary
Standards/ ▼Strands													
Concepts and Procedures													
1. Number Sense													
2. Measurement													
3. Geometry													
4. Prob. and Stats													
5. Algebra													
Problem Solving													
1. Investigate													
2. Formulate and Define													
3. Solutions													
Reasoning													
1. Analysis													
2. Prediction													
3. Conclusion/ Verification													
Communication													
1. Gathering													
2. Organizing/ Interpreting													
3. Represent/ Share													
Connections													
1. Within math													
2. To other disciplines													
3. To real-life situations													
Comments:													
										Overall Grade			

* Essential Academic Learning Requirements

> "The use of columns in a grade book to represent standards instead of assignments, tests, and activities is a major shift in thinking for teachers."
> —Marzano & Kendall, 1996, p. 150

Marzano (2006) further suggests that in choosing what to include in each measurement, topic teachers should consider the concepts of *unidimensionality*—"a single score on a test represents a single trait or dimension that has been assessed" (p. 14)—and *covariance*—as ability in one dimension increases, so does that in another. As much as possible, measurement topics should be unidimensional, and if more than one learning goal is included, they should be covariant. Using these concepts, Marzano provides a list of measurement topics he suggests for use in kindergarten through Grade 10; the list has 12 topics for language arts, 18 for mathematics, and 13 each for science and social studies (p. 24).

Another approach to organizing learning goals is offered by Guskey (2005). He suggests that teachers develop "tables of specification" based on the following categories: knowledge of terms, knowledge of facts, knowledge of rules and principles, knowledge of processes and procedures, ability to make translations, ability to make applications, and skill in analyzing and synthesizing.

Another aspect of organizing learning goals that requires consideration is the nature of the learning goals. In the Washington mathematics example, the learning goals were classified as skills (e.g., problem solving) or discrete knowledge categories (e.g., measurement). This is the most common approach. It may, however, be better to take a more conceptual approach, as it then becomes clear that learning is a process that occurs over time. Curriculum, instruction, and assessment can be clearly focused on students developing understanding of the concepts, and these concepts become the base for grading and reporting. An example of this approach can be found in AERO (American Education Reaches Out) social studies standards developed for use in international schools. The standards/strands are Time, Continuity, and Change; Connections and Conflict; People, Place, and Environment; Culture; Society and Identity; Governance and Citizenship; Production, Distribution, and Consumption; and Science, Technology, and Society (AERO, n.d.).

A somewhat different approach can be found in Ontario, Canada (Figure 1.6). For all elementary and secondary subjects, achievement charts (see an example in Chapter 2, Figure 2.8) have been developed that provide the performance standards. Each chart has descriptors of four levels of achievement for each of four categories of knowledge and skill. These categories are now consistent across all subjects in Grades 1 to 8 and the 15 disciplines included in the secondary curriculum.

This approach was developed to

- "provide a common framework that encompasses all curriculum expectations for all grades and all subjects/disciplines;
- guide the development of assessment tasks and tools (including rubrics);
- help teachers to plan instruction for learning;
- assist teachers in providing meaningful feedback to students;
- provide a variety of aspects (e.g., use of thinking skills, ability to apply knowledge) on which to assess and evaluate student learning." (Ontario Ministry of Education, 2004, p. 2)

Figure 1.6 Summary of Evidence for Ontario Secondary Subjects or Courses

Student:_____

ACHIEVEMENT EVIDENCE

Assessments / Knowledge/ Skill Categories													Summary
Knowledge/ Understanding													
Thinking/Inquiry													
Communication													
Application													
Comments													
								Most consistent level of achievement with consideration for more recent					

Expectations (as learning goals are called in Ontario) have to be classified into the most appropriate category, and then an approach identical to that for Washington mathematics can be followed (see Figure 1.6). Further examples of similar approaches can be found in Reeves (2000, p. 13), Marzano (2000, pp. 106–118), and Cooper (2006, pp. 391–394).

One perceived drawback of this approach is some additional bookkeeping, but, especially if rubrics are used, it takes less time to score assessments. Therefore, this time savings can be balanced against the additional time needed to record scores. In addition, while every strand or category should be assessed enough times over a year or course to make valid and reliable judgments of achievement, it is not essential that there be scores or grades for each strand/category in each grading period.

If overall grades are required, another aspect of this approach that teachers need to consider is whether each strand or category is of equal significance or whether some strands or categories are more important for the whole year or for any particular grading period. It is usually best to start from the position that each strand or category is of equal significance and later make appropriate adjustments if it is obvious from the emphasis in the curriculum policy or in the way the subject is taught that one or more strands or categories are of greater significance than others. An example of this uneven distribution is the recommendation from the provincial association for physical and health education (PHE) teachers in Ontario that the application category be assigned a weight of 60–65 percent for Grade 9 and 10 PHE.

If grading plans are approached in this fashion, the learning goals become the set, and the assessment methods become the subset. As has been shown, teachers identify for each assessment what components (or questions) fit with what learning goals and then record separate scores for each.

Grading is an exercise in professional judgment, not just a mechanical, numerical exercise.

A detailed examination of this approach to grading is provided by Marzano and Kendall (1996), who say that

> first and foremost, the teacher must stop thinking in terms of assignments, tests and activities to which points are assigned, and start thinking in terms of levels of performance in the declarative and procedural knowledge specific to her subject area. (p. 147)

They also acknowledge that "the use of columns in a grade book to represent standards, instead of assignments, tests, and activities, is a major shift in thinking for teachers" (p. 150). This illustrates clearly that one of the major challenges for teachers in implementing standards-based grading is how to organize the grade book. The one-page-per-student approach is illustrated in Figures 1.2 through 1.6. For teachers who find that having a separate page for each student is impractical, an example of a single spreadsheet for each class can be found in Chapter 9, Figure 9.4b.

One subject in which this approach has been seen to be particularly difficult is physical education. Melograno (2007) provides helpful direction

when he suggests that teachers of this subject should grade and report student achievement based on the standards developed by the U.S. National Association for Sport and Physical Education (Competency in Motor Skills and Movement Patterns; Understanding of Movement Concepts, Principles, Strategies, and Tactics; Regular Participation in Physical Activity; Health-Enhancing Level of Physical Fitness; Responsible Personal and Social Behavior; and Values Physical Activity) and examples of how this could be done.

Related issues raised by Marzano and Kendall (1996) are how final grades are determined and how student performances are reported. The first issue is dealt with in Chapter 9 of this book and the latter issue in Chapter 11. It is sufficient to note here with regard to reporting that if grades are related to learning goals, it is at the very least highly desirable, if not essential, that report cards provide opportunities for teachers to communicate specific information on each learning goal in addition to an overall grade.

Learning Goals and Passing Grades

A final issue that needs to be considered in connection with this guideline is whether students should receive credit for a course if they have not demonstrated mastery of the critical learning goals. In the example in Figure 1.2, there are five strands in Washington mathematics. Although unlikely, it would be possible for a student to obtain very low scores on two strands while obtaining sufficiently high scores on the other strands to obtain a passing overall grade. Teachers and schools need to decide if this is acceptable. If they really believe all or some strands are critical, then students should not be able to obtain credit unless they have achieved a reasonable level of competence in—ideally mastery of—each of those strands. This is certainly the approach taken with pilots (and plumbers): a student pilot has to be competent in takeoffs, flying the plane in the air, landings, and navigation before receiving certification. It is not acceptable that the pilot be excellent in three of these and less than competent in the fourth, but this is the performance we reward in high schools all the time with traditional grading.

Some schools are moving toward a true standards-based approach. One such school is the Foxcroft Academy in Maine; it requires students to demonstrate competency in most standards and partial competency in the remaining standards in each subject to receive credit. For example, in English, competency is required in 6 of 8 standards with partial competency in the remaining 2, math requires competency in 9 of 11, history 10 of 13, and physical education 2 of 3. In health, 3 standards are specified as requiring competency (Health Information, Health Promotion, and Decision Making), with the other standards requiring partial competency. A slightly different approach is taken at Poland Regional High School, also in Maine, where it requires competency in 17 courses and credit in at least seven other subjects for graduation. Yet

another approach has been developed at Eagle Rock School in Colorado (see Easton, 2007).

If teachers use this approach, it obviously complicates the grading process, but it supports the concept that grading is an exercise in professional judgment, not just a mechanical, numerical exercise.

It also illustrates the interconnectedness of the grading guidelines because Guidelines 4 and 5 (see Chapters 4 and 5) become absolutely critical. Formative assessment has to be used to provide information to students and to teachers about progress (Guideline 4), and students need to have growth acknowledged appropriately and have varied opportunities to demonstrate competence (Guideline 5).

WHAT'S THE BOTTOM LINE?

- Teachers should base grades on learning goals (standards, expectations, outcomes, etc.), not assessment methods.
- Teachers should have a clear understanding of what learning results are expected.
- Reporting should allow for focus on information on each learning goal.
- Credit should be granted only when students have mastered the critical learning goals. (Mastery is the ideal.)

This guideline has these practical implications:

- Teachers use grade books where the columns primarily represent the learning goals and secondarily represent assessment methods.
- Report cards at all levels provide grades on each learning goal and only provide an overall subject grade if necessary. I believe that single-subject grades are necessary only for Grades 11 and 12. This is to meet the needs of colleges and universities, not for any pressing educational or communication reason.

WHAT'S MY THINKING NOW?

Guideline 1: Relate grading procedures to learning goals (i.e., standards).

Analyze Guideline 1 for grading by focusing on three questions:

Why use it?

Why not use it?

Are there points of uncertainty?

After careful thought about these points, answer these two questions:

Would I use Guideline 1 now?

Do I agree or disagree with Guideline 1, or am I unsure at this time?

See the following for one person's reflections on Guideline 1.

A REFLECTION ON GUIDELINE 1

Why use it?

- Links to basis for curriculum, instruction, and assessment.
- Realistically reflects intentions of course/grade.
- Provides clear goal/focus.
- Students know why they received grades.
- Consistency and fairness.
- Makes teachers accountable.

Why not use it?

- Enforces loss of creativity.
- Requires too great a shift in thinking/practice.
- Learning goals are often vague.
- May engender negative community reaction.
- Requires a large amount of work to select learning goals, develop grading plan, etc.

Points of Uncertainty

- Clarity of learning goals?
- Fair to all learning styles?
- Weighting learning goals?
- Mastery or pass/fail?
- Specificity of learning goals?
- How many learning goals?