OVER THE LAST TEN TO FIFTEEN YEARS, each province and territory in Canada has introduced curriculum based on clearly defined learning goals. They use different labels – for example, “learning outcomes” in Manitoba and Alberta, “expectations” in Ontario – but all are intended to be “standards-based” systems where the focus is on outputs (“students will”) rather than inputs (“students will be provided with opportunities to”). At the same time, there has been a growing focus on the key role of assessment in the learning process, and each jurisdiction in Canada has put increasing emphasis on principles of quality assessment and the need for assessment-literate communities. This focus is clear, for example, in *Rethinking Classroom Assessment with Purpose in Mind*, published in 2006 by Manitoba Education, Citizenship and Youth, in collaboration with the Western and Northern Canadian Protocol for Collaboration in Education.

These two parallel developments have led to a growing recognition of a mismatch between many traditional practices and the requirements of standards-based, assessment-literate systems. Practices such as combining achievement and behaviour in grades, the use of penalties for “late work”, the use of zeros as punishments, and the role of homework in grading have all come into question. Some of the provincial policies designed to address these concerns have been controversial.

The purpose of this article is to provide some background to these issues by providing information about the history of grading, the purposes of grades, the impact of grades, and how grades have and should be determined.

A SHORT HISTORY OF GRADING

Although student assessment and reporting on student achievement has been part of education for centuries, the use of grades is a fairly recent development. Prior to about 1880, reporting was in a narrative format and often simply listed the skills and concepts that each student had mastered, but by the late 1800s and early 1900s schools started to use letter or number symbols to summarize student learning. This reductionist movement began in universities, and then moved to K-12 schools, especially to high schools, in response to a growing student population.

By 1910 the use of percentage grades was common and widely accepted, especially in high schools. This approach, however, came into question with the publication in 1912 and 1913 of research studies by Starch and Elliot that showed the unreliability of teacher-marking using percentages, first in English and then in geometry. As a reaction to this research many schools turned to grading scales with three to five categories, and the five-level A-F scale had become a common approach by the early 1920s.

DEFINITION OF TERMS

For the purposes of this article, a **grade** is defined as a symbol (letter or number) on a report card that summarizes student achievement. A **mark** or **score** is defined as the symbol (letter or number) given to any student test or performance that provides evidence of student achievement.

Commonly in Canada the term “mark” is used for both of these, but as there are two different processes involved, the meaning of each is clearer when the two terms are used as defined above.

WHEN, WHY, WHAT IMPACT, AND HOW?

Grades are a tool used to assess student learning and progress. They are used to communicate information about student achievement to students, parents, teachers, and other stakeholders. The purpose of grades is to provide an objective measure of student performance, to encourage student effort, and to facilitate the planning of future instruction. Grades can also be used to identify students who may need additional support or intervention.

Grades can have a significant impact on students, teachers, and schools. They can affect student motivation, self-esteem, and confidence. They can also influence teacher decision-making, curriculum planning, and school policy. However, grades are not without their limitations. They can be influenced by a variety of factors, including the difficulty of the curriculum, the quality of instruction, and the student’s prior knowledge and skills. They can also be used inappropriately or misinterpreted, leading to unfair or inaccurate assessments.

In order to be effective, grades must be meaningful and transparent. They must be based on valid and reliable assessments of student learning, and they must be communicated in a clear and understandable manner. They must also be used in conjunction with other forms of assessment, such as formative assessments and feedback, to provide a complete picture of student progress.

Grades are an important tool in education, but they must be used appropriately and with care. They can be a valuable source of information for students, teachers, and schools, but they must be used in a way that supports student learning and development.

KEN O’CONNOR

1. *Rethinking Classroom Assessment with Purpose in Mind*, published in 2006 by Manitoba Education, Citizenship and Youth, in collaboration with the Western and Northern Canadian Protocol for Collaboration in Education.

2. By 1910 the use of percentage grades was common and widely accepted, especially in high schools. This approach, however, came into question with the publication in 1912 and 1913 of research studies by Starch and Elliot that showed the unreliability of teacher-marking using percentages, first in English and then in geometry. As a reaction to this research many schools turned to grading scales with three to five categories, and the five-level A-F scale had become a common approach by the early 1920s. But soon concerns began to be voiced about the subjective nature of this type of grading, and so in the 1930s grading on the curve became increasingly popular, reflecting the common belief that student abilities were distributed along a normal curve.
THE PURPOSE OF GRADES

Trumbull suggests that there have been three main purposes for grading: giving feedback, motivating, and sorting. She expands these broad purposes as illustrated in figure 1.1

<table>
<thead>
<tr>
<th>Giving Feedback</th>
<th>Motivating</th>
<th>Sorting</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Inform parents (and students)</td>
<td>• Encourage students to improve or keep working</td>
<td>• Make placement or grouping decisions</td>
</tr>
<tr>
<td>• Account to community</td>
<td>(promote student learning)</td>
<td>• Certify competence, permit graduation</td>
</tr>
<tr>
<td>• Recognize good work</td>
<td>• Reward students who are doing well</td>
<td>advance student to next grade</td>
</tr>
<tr>
<td>• Identify unacceptable work</td>
<td></td>
<td>• Predict future achievement</td>
</tr>
<tr>
<td>• Promote student self-evaluation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Identify instructional gaps</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

For students motivated by grades, their main impact has been to turn school into a grading “game” not a learning “game.”

Clarity of purpose is critical to everything we do; it is our compass and provides us with a sense of direction, especially when there are questions about how we should proceed. Grades do not serve all purposes equally well, so Trumbull’s list of purposes creates a serious problem: with so many purposes for grades, some of which conflict (sorting often conflicts with feedback), it is difficult to get a clear sense of direction. Thus, we need to prioritize.

Although provincial ministries of education generally do not state the purpose of grades, it is reasonable to assume that if they did, they would agree with Brookhart and Bailey and McTighe; the primary purpose of classroom assessment, according to most ministries, is to gather information that informs teaching and learning. However, the critical point here is that ministries, schools, and school districts must clearly identify their primary purpose in grading – whatever it may be – so that subsequent decisions can flow from that purpose.

THE IMPACT OF GRADES

Grades are an efficient way to summarize student achievement and have traditionally been believed to motivate students to work hard and behave well. This has been the case for students who receive the grades that they expect or believe they deserve, but for students who receive grades lower than they expect or believe they deserve, grades have often been de-motivators.

For students motivated by grades, their main impact has been to turn school into a grading “game” not a learning “game” because the student’s focus becomes the accumulation of points, not learning. Guskey and Bailey note, “The currency of points dominates the academic economy of classrooms...Savvy students keep track of current exchange rates, calculating far in advance the exact number of points they need to obtain the grade they want, and adjust their efforts accordingly. They know they must plan cautiously since they can lose points or be fined for certain transgressions...They also make note of contingencies that allow them to earn extra points or bonuses.”

For students not motivated by grades, the grade-driven economy of schools has caused them to withdraw from learning, frequently becoming behaviour problems and/or dropping out.
Kohn states that “researchers have found three consistent effects of using – and especially, emphasizing the importance of – letter or number grades”:  

- Grades reduce students’ interest in learning; 
- Grades reduce students’ choice for challenging tasks; 
- Grades reduce the quality of students’ thinking.8

The basic question raised by the motivational impact of grades is – or should be: What type of motivational environment should we have in place in schools? Grades have contributed to an environment that maximizes extrinsic motivation, but most schools now say that they are trying to develop students into lifelong learners – a goal that is not achieved through extrinsic motivation.

**HOW GRADES ARE DETERMINED**

For most of the history of grades, the mission of schools was seen as sorting students into a reliable rank order so that they could be appropriately placed into educational programs and the world beyond school. But over the last twenty years the mission of schools has changed; it is now to ensure that all students have met essential learning goals. This shift has led us to ask whether the new mission calls for a new way of determining grades. Issues that need to be examined include the reference points used to determine grades, the purpose and use of assessments, the quality of assessments, the basis for grades (both the “what” and the “how well”), the ingredients included in grades, and the appropriateness (or lack thereof) of the mean as a summary of student achievement.

**Reference Points**

Traditionally a norm-referenced interpretation has been placed on assessment results, and this is still appropriate for standardized testing; but for classroom assessment we must now use a criterion-referenced approach. Classroom assessment expert Rick Stiggins puts it this way: “We have emerged from an era of comparing students with other students based on achievement to a time when we compare student performance to pre-set standards; and now we ask who has and who has not met the standards.”9

**Purpose and Use of Assessments**

Until recently everything students did, regardless of purpose, has been included in grades; but as our understanding of the various purposes of assessment has improved, we have come to understand that a distinction should be made between assessment of learning (summative assessment) and assessment for learning (formative assessment). Summative assessments should provide the evidence used to determine grades, while formative assessments support learning by providing students with descriptive feedback that they can use to improve. This has significant implications for classroom practice and is supported by a large and growing body of research demonstrating that when formative assessment is done well, subsequent student achievement improves dramatically.10

**The Quality of Assessment**

When the primary concern was the rank order of students, the quality of assessment was not a major issue; but when our primary concern is the competence of all students, it is essential that all assessments provide quality evidence of student achievement. Thus it is now essential that all teachers are assessment literate and understand that quality assessments require clear learning goals, clear purpose, and sound design. The latter requires target-method match and assessments that are well written, well sampled, and free of bias or distortion.

If the focus is on learning goals, it is essential that grades be as pure measures of achievement as possible.

**The Basis for Grades**

The **What.** Traditionally the basis for grades has been assessment methods or activities, and the categories in teachers’ grade books have been, for example, tests, projects, and assignments. In systems based on learning goals, the basis should be the learning goals themselves, so the categories in an English teacher’s grade book, for example, should be reading, writing, listening, language, and literature.

The **How Well.** This has most commonly been based on an accumulation of points with a set average percentage determining pass or fail. As Canadian assessment expert Damian Cooper notes, “Most Canadian provinces use 50 percent as their pass/fail cut-point. Knowing 50 percent of the material taught can hardly be considered ‘proficient’. Nobody wants to fly with a pilot who scored 50 percent on his or her exams in flight training school. Pass/fail cut-points are an outdated relic of norm-referenced approach to grading.”11 It has become increasingly obvious that the percentage system is incompatible with a learning goals system and that what is needed instead is clear descriptions of a limited number of levels.12

**Ingredients included in Grades**

Traditionally grades have resulted from a rather uncertain mix of achievement and behaviour (attendance, punctuality, following rules, etc), and so it has often been difficult to tell what a grade means. If the focus is on learning goals, it is essential that grades be as pure measures of achievement as possible without penalties for such behaviours as handing assessment evidence in late. Ideally, the behaviours we value are reported separately on the report card.13

**The Appropriateness of the Mean as THE Measure**

Until recently, virtually the only calculation used to determine grades has been the mean, but we teach in math that the mean is an inappropriate measure of central tendency when there are outlier scores. The outlier scores that students usually have are low outliers (especially zeros) so there has been a growing realization that consideration should be given to the median or mode as a more accurate summary of student achievement. There has also been increasing recognition that no measure of central tendency adequately allows for appropriate representation of more recent achievement, so that ultimately grading is (or should be) seen not as just a numerical, mechanical exercise but as an exercise in professional judgement.
One result of all these considerations has been the understanding that grading must move from the individual, idiosyncratic, private practice it has been traditionally to a shared practice based on agreed upon principles or guidelines. Almost all of the experts quoted in this article and other assessment specialists have developed guidelines over the last few years, and while there are differences in emphasis and wording, a consensus about what is best practice in grading has emerged.\(^\text{14}\)

**CONCLUSION**

Brookhart states, “In a perfect world there would be no grades – at least, not as we know them now.”\(^\text{15}\) This is probably a reasonable statement, but provincial policies ensure that, for the foreseeable future, there will be grades in Canadian schools. To make grades more educative, provincial ministries of education must state clearly what they see as the primary purpose of grades and then ensure that their policies are aligned with that purpose. When that happens it will be possible to say that we have learned from the history of grading and that schools are focused on learning, not grades.\(^\text{1}\)


**Notes**

1. The information in this section has mostly been obtained from T.R. Guskey and J.M. Bailey, Developing Grading and Reporting Systems for Student Learning (Thousand Oaks, CA: Corwin, 2001).
3. E. Trumbull, “Why Do We Grade – And Should We?” in Trumbull and Farr, 24.
4. Trumbull, Figure 2.1, 25.
7. Guskey and Bailey, 19.
12. The two most highly regarded high school programs in the world – Advanced Placement and International Baccalaureate use only levels – AP 5 and IB 7.
13. See for example the Ontario Provincial Report Cards with nine “learning skills” on the elementary report card and five on the high school report card. Starting with the 2010-11 school year the same six learning skills will be on all report cards. These report cards can be found at www.edu.gov.on.ca/eng/document/forms/report/1998/report98. html (accessed December 28, 2009).
15. Brookhart, 2.