Grades and Grading Practices

William Strein
Nicole Meshebesh
University of Maryland, College Park

Grading and reporting of student achievement and progress are universal features of formal educational environments from kindergarten to graduate school. Many teachers believe that grades are overemphasized, but it is recognized that grades have important consequences for students’ present lives (rankings, awards, failure), for postsecondary educational opportunities, and for employment (McMillan, Myran, & Workman, 2002). Grading issues are extremely controversial, with little consensus among stakeholder groups (e.g., parents, teachers, students, administrators, and community members), to the extent that efforts to make even moderate revisions in grading and reporting procedures are extraordinarily difficult (Guskey & Bailey, 2001). This chapter explores key issues in grading and reporting, and provides information from the literature that will assist educators, and those who consult with them, in improving grading and reporting practices for K–12 students in both general and special education programs.

BACKGROUND AND DEVELOPMENT

Functions and Purposes of Grading

Grading typically serves four basic purposes: (a) certifying that students have achieved a specific level of accomplishment or mastery (e.g., high school graduation), (b) selecting students for some educational or occupational path (e.g., college admission, or employment), (c) informing students and parents about the student’s progress, and (d) motivating students (Natriello, 1992). Grading policies and practices that serve one purpose particularly well may actually hinder another purpose (Guskey & Bailey, 2001). For example, certification and selection are often well served by high standards and competitive grading practices, but when grading is used for these purposes, it tends to be most motivating to already high-achieving students, and may actually decrease motivation of lower achieving students (Natriello, 1987). As long as a teacher is using a single-indicator reporting system (e.g., one letter grade or percentage number for English), there seems to be no reasonable resolution to the problem of trying to ascertain clearly the meanings of the grade assigned.

Although assessment and grading are inextricably intertwined, it is helpful to distinguish between them. Assessment refers to the objective measurement of a student’s learning. Grading refers to the teacher’s evaluation of the adequacy of a student’s learning.

Typical Approaches to Grading

This section presents common approaches to grading (and discusses the advantages and disadvantages of each), research on what grading practices teachers actually use, and what factors affect teachers’ grading practices. The majority of practices that are typically used are based on four broad approaches to grading: letter grades, other categorical systems, percentage grades, and standards-based grades.

Letter grades (ABCDF) remain the most commonly used system in secondary schools; they are less frequently used in elementary schools. Many experts (e.g., Stiggins, 2005) consider traditional grading to be inevitable, at least for the immediate future. One of the clearest advantages to traditional letter grading is that it is familiar to parents, teachers, and students alike. Letter grades also provide a succinct description of achievement and performance. Despite some reservations, teachers find letter grades to be generally helpful (Bursuck et al., 1996),
Children’s Needs III

at least for nondisabled students. Disadvantages of letter grades include (a) inappropriate collapsing of different information into one symbol, (b) arbitrary grade cutoffs that are difficult to make, and (c) lack of comprehensiveness of other methods, such as standards-based grading (Guskey & Bailey, 2001).

Other categorical systems of measurement are beginning to be used in place of letter grades. For example, the state of Kentucky (as reported in Guskey & Bailey, 2001) adopted a standards-based system that uses the categories of Distinguished, Proficient, Apprentice, and Novice. Advocates of such systems claim that they are more affirming of the student and are more descriptive than letter grades. Disadvantages are similar to those of letter grades. Also, there is little evidence as to whether such systems truly have fewer negative connotations (Guskey & Bailey, 2001). In addition, such systems may imply a carefully constructed standards-based system (see below), when in fact the categories are simply letter grades that are relabeled.

Percentage grades are second in popularity only to letter grades and have a long history of use in American schools. This system is typically more popular in middle and high schools. The strongest advantage of percentage grades is that they have the potential of providing large discrimination in evaluation of achievement. Again, a disadvantage includes collapsing of much information into a single indicator. For example, averaging scores on three quizzes and a major exam to produce an overall percentage may not give enough weight to the more comprehensive indicator of knowledge (Guskey & Bailey, 2001).

Standards-based grading (Guskey & Bailey, 2001) involves setting clear criterion-referenced standards in any given subject and then reporting levels of mastery in reference to each standard. A criterion-referenced approach bases grades on students’ individual attainment of some previously defined levels of achievement, and it involves no comparison to other students. Most applications are currently geared to the elementary level. For example, a standards-based reporting system in elementary school math might include items such as “Subtracts single-digit numerals,” “Subtracts two-digit numerals without regrouping,” or “Subtracts two-digit numerals requiring regrouping.” Each item might then be evaluated with a notational system such as “1 – Beginning Standard,” “2 – Approaches Standard,” “3 – Meets Standard,” “4 – Exceptional: Exceeds Standard” (Guskey & Bailey, 2001, p. 88). Such systems may summarize performance within an academic area with a notational system similar to an overall “grade” in math. Some systems dispense with summarized grades altogether.

Advantages of standards-based grading include providing important clear information about students’ performance, and providing information that may be useful for both diagnostic and prescriptive purposes. Two disadvantages are that (a) such systems are labor-intensive, especially in terms of initially developing and implementing the system, and (b) reporting systems with large numbers of standards and corresponding indicators are often too complicated for parents to understand.

Teachers’ actual practices. Several recent large surveys (Bursuck et al., 1996; McMillan, 2001; McMillan et al., 2002) and a comprehensive review (Brookhart, 1994) of earlier research provide a clear picture of the extent to which teachers use the above practices of grading. Surveys of elementary and secondary school teachers produce generally consistent results. Additionally, the findings from six newer (1996–2002) studies are very consistent with Brookhart’s conclusions arising from her review of 19 studies published from 1985 to 1993. Apparently, teachers’ grading practices have been quite stable over at least two decades.

In principle, three general categories cover most criteria for grading. Product criteria are based on specific student achievement or levels of performance. Process criteria refer to how a student performs and not what the student has attained, for example, effort, work habits, and perhaps performance on daily assignments. Progress criteria emphasize student gains (or losses) over time and are highly individualized across students. One of the most enduring findings of research on teachers grading practices is that at both the elementary and secondary school levels and across content areas, teachers nearly universally use hodgepodge grading, a practice that is supported widely by both students and teachers (Cross & Frary, 1999; McMillan et al., 2002). Hodgepodge grading refers to combining product, process, and progress criteria into a single grade. There is great variation as to how individual teachers weight such factors (Brookhart, 1994; McMillan, 2001).

Notwithstanding common perceptions, disruptive student behavior appears to contribute little to grading (McMillan et al., 2002). Although teachers routinely include nonachievement factors, sometimes referred to as academic enablers, such as effort, participation, and responsibility (Cross & Frary, 1999; McMillan et al., 2002), academic performance is consistently the most important factor in grading (McMillan, 2001; McMillan et al., 2002).
Parents and many educators stress the importance of considering academic enablers when grading low-achieving students, students who recently have made good progress relative to their own previous baseline, or students who are persisting despite some difficult personal obstacles. However, the increased focus on meeting standards has decreased the acceptability and use of nonachievement-based indicators in the grading process.

In elementary school, grading practices vary little across content areas (e.g., language arts and math), although increased importance is placed on homework and assessments in the upper elementary grades (McMillan et al., 2002). Similar to their elementary school counterparts, secondary teachers use a variety of factors to guide their grading practices. Most notable among those are that (a) academic achievement is the most important factor, and (b) nonachievement factors are not always used in determining grades, whereas this practice is almost universal in elementary school (McMillan, 2001). Variation also occurs across content areas at the secondary level. For example, in surveying the grading practices of secondary teachers (Grades 6–12), McMillan found that social studies teachers used effort and participation and extra credit more often than math teachers. Similarly, high school science teachers used effort in only a minor way and rarely included other nonachievement factors (Feldman, Alibrandi, & Kropf, 1998).

At all grade levels, tests, quizzes, and in-class homework accounted for the majority of the basis for assigning grades, with tests and quizzes becoming increasingly emphasized at the secondary level (Bursuck et al., 1996). Out-of-class homework and class participation had moderate to strong effects on grades (Cross & Frary, 1999), with an increased importance placed on homework at later grades (McMillan et al., 2002). In terms of reference points for anchoring grades, Feldman et al. (1998) found that 50% of high school science teachers used criterion-referenced grading, 28% graded students relative to one another, 16% based grades on the student’s individual ability, and 2% assigned grades based on the student’s academic growth from some previous point.

Given the substantial variation in teachers’ grading practices, it is useful to ask what factors influence teachers’ decisions about grading practices. A substantial body of evidence indicates that teachers’ philosophies and beliefs about education, learning, and instruction constitute the single greatest influence on teachers’ grading practices. For example, in a large survey of the assessment and grading practices of teachers in Grades 3–5, McMillan et al. (2002) found no differences in grading practices across grade levels. Importantly, however, they found that variability in teachers’ grading practices within schools was greater than the variability between schools. Using these results and other published research, which reported large and unpredictable between-teacher variability (e.g., Cizek, Fitzgerald, & Rchor, 1995), McMillan et al. concluded that, although school policies and other external factors have some effect, “individual teacher preferences are more important than are differences between schools in determining grading practices . . . [and] school and student characteristics as a whole are less important than are individual beliefs” (p. 212). Similarly, in an interview study of 24 teachers of Grades 5–12, McMillan and Nash (2000) concluded that “[grading practices] were influenced most heavily by internal beliefs and values that were frequently idiosyncratic, reflecting teachers’ philosophies of education” (p. 6).

Notwithstanding the overriding importance of teachers’ beliefs, external influences such as school policies on grading (Christiansen & Vogel, 1998), subject areas and academic level being taught (e.g., regular vs. advanced placement classes; McMillan, 2001), and assessment systems being used (McMillan & Nash, 2000) influence teachers’ grading practices. Easily available computer technology (e.g., special grade-book software) appears to have affected grading by fostering such practices as combining grades from various activities to form weighted composite grades (Feldman et al., 1998). Without question, the standards movement and the No Child Left Behind Act (NCLB) exert increasingly strong influence on teachers’ grading practices (see later sections of this chapter).

**PROBLEMS AND IMPLICATIONS**

**Relationship Between Assessment and Grading**

Assessment and grading practices are interconnected but differentiable aspects of teaching. Greater attention and emphasis have been placed on classroom grading practices within the professional literature, whereas research pertaining to assessment has focused on standardized or large-scale testing as opposed to teacher–classroom assessment practices (McMillan & Workman, 1998). Classroom assessments are an integral part of instruction. The results often provide invaluable information about
students' understanding of content that can guide teacher instruction. However, experts question teachers' competence in designing and evaluating classroom assessments that are consistent with recommended measurement principles. Teachers' limited assessment competencies can affect not only the quality and effectiveness of the assessment measure, but also have consequences for students' learning, their overall motivation, and their perceptions of their ability level (McMillan & Workman, 1998).

In an interview study of teacher classroom assessment and grading practices across Grades 5–12, McMillan and Nash (2000) found that teachers use a variety of assessments in order to balance external pressures and personal beliefs. For example, the majority of the teachers interviewed indicated that formative and more informal assessments (e.g., daily quizzes, observation, daily checks of student understanding), as well as constructed response assessments (e.g., open-ended essay and short-answer questions, and presentations) were very useful and provided the most valuable information regarding student performance. However, objective measures (e.g., selected response assessments) also were often used and incorporated into teachers' repertoire of assessments. The advantage of objective measures is that they do not possess the ambiguity and subjectivity inherent in evaluating constructed response assessments. Thus, using objective measures can alleviate conflicts with parents who may question or challenge the grade on constructed response assessments.

At the elementary school level, objective tests are used most often, but many teachers also use constructed response assessments. Also, elementary teachers distinguish between levels of assessment that require rote recall versus higher order thinking (McMillan et al., 2002). At the secondary level, English teachers use constructed response assessments more than math, social studies, and science teachers. Secondary teachers of above-average and advanced placement (AP) classes tend to assess higher order thinking skills, whereas teachers of average level or below-average classes tend to assess recall knowledge (McMillan, 2001).

**Alternative Assessments**

Increased emphasis on alternative assessments that focus on higher order thinking and problem-solving skills has greatly influenced recommendations for classroom assessments. This emphasis represents a shift from traditional assessments, such as selected response and other objective measures. Generally, the advantage of using alternative assessments is that they “require the active construction of meaning rather than passive regurgitation of isolated facts” (McMillan & Workman, 1998, p. 2). The caveats, for schools that want teachers to implement alternative assessments, are that many teachers may need inservice training on using such assessments. Also, more frequent communication with parents may be needed in order to clarify the expectations for students' learning and how students will be evaluated using the alternative assessments (Carlson, 2003).

Alternative assessments include authentic assessments, portfolios, journals, grading rubrics, and individual and group projects (McMillan & Workman, 1998). Authentic assessments link the student evaluation to real-world situations. Portfolios include a collection of completed work, either electronic or hard copy. The benefits to using portfolios are that they offer a means of evaluating student performance over a period of time, as well as provide the opportunity for students to become actively engaged in the development, design, and content within the portfolio. Grading rubrics delineate the criteria by which student performance will be evaluated. The benefits of using grading rubrics are that they inform students of the teacher's expectations prior to beginning an assignment, which can help guide students as they complete the assignments, and rubrics provide a means to communicate clearly to parents how the student's grade on the assignment is derived (Jackson & Larkin, 2002). Individual or group projects are desirable because they foster active student participation. The primary advantages of using group projects are not only that they actively engage students, but that they also require cooperation among members of the group to jointly produce the project. One of the significant disadvantages associated with group projects is determining how to grade the participants so the grading is fair and so it accurately reflects the contributions of each student.

**Real-World Considerations Versus Expert Recommendations**

A consistent finding in the research literature on grading practices is that teachers often do not follow established measurement principles, resulting in incongruence between the recommendations of experts in measurement and the actual practices of teachers (Brookhart, 1994). Measurement experts argue that basing grading on sound measurement principles would
eliminate the subjectivity of grades that are based on the "hodgepodge of factors" and would also result in increased consistency of grading practices across teachers (McMillan et al., 2002). Experts recommend the following practices (e.g., Cross & Frary, 1999; Guskey & Bailey, 2001):

- Use product-oriented rather than process- or progress-oriented criteria. For example:
  - Base grading *exclusively* on measures of current achievement.
  - Eliminate "academic enablers" (grading for effort, participation, and responsibility) in grading.

- Base grading on the evidence that best depicts student achievement at the point at which the grade is assigned. For example:
  - Give greater weight for more recent evidence.
  - Give greater weight to the most comprehensive sources of information, e.g., comprehensive tests versus daily work.
  - Rank order evidence in terms of importance to learning.
  - Eliminate the assigning of zeros for missing work; use an Incomplete (I) instead.

- Eliminate the common practice of "hodgepodge grading," such as combining effort, achievement, ability, and so forth into a single grade.

- Inform students in advance about expectations and criteria used for grading (this being the only recommendation that teachers seem to adhere to).

Teachers' reluctance to follow the measurement experts' recommendation—to assign grades based solely on achievement—is likely related to teachers' understanding that grades serve other functions in addition to indicating students' level of achievement. It is also likely that the achievement-only approach to grading de-emphasizes the function of grades as a pedagogical tool, that is, "... [promoting] student success in general in many areas important to schooling, including both academic and [nonacademic] achievement [factors] such as responsibility, effort, improvement, participation, and cooperation" (McMillan et al., 2002, p. 212). Moreover, teacher grading practices are likely to be affected by taking into account individual student differences, wanting to encourage student involvement and motivate students, wanting students to succeed, and wanting to avoid too many failures (McMillan, 2001).

Many teachers lack training in recommended assessment and grading practices (Salend & Duhaney, 2002), but the literature is mixed as to whether such training actually increases teachers' use of recommended practices (Bursuck et al., 1996; Cross & Frary, 1999). Although teachers typically work in isolation, they do not work in a vacuum that is free of external pressures, demands, and challenges associated with teaching a class of diverse learners. Changes in existing grading practices are unlikely to occur unless measurement experts work collaboratively with educators to determine best practices for assessment and grading that both incorporate sound measurement principles and take into account daily classroom functioning (Brookhart, 1994).

**The Standards Movement and the No Child Left Behind Act: Implications for Grading Practices**

At the beginning of the 21st century, there is probably nothing more influential or more controversial in American education than the increased emphasis on accountability standards. The history of and controversies surrounding the standards movement are far too extensive and complex to even summarize here (also see chapter 46, "Grade Retention and Promotion"). Although the details vary, the signature feature of the standards movement is the use of student evaluation procedures external to the classroom. That is, the standards and accompanying evaluation tools used are not designed by the classroom teacher. In addition, the standards movement promotes the use of statewide or even national criteria and evaluation methods for assessing the adequacy of student progress. This is in sharp contrast to traditional grading practices in which the evaluation standards are set, de facto, by the teacher, even when school district grading policies and guidelines are in place.

The No Child Left Behind Act of 2001 does not address the issue of grading per se. However, under NCLB all public school children must be tested annually on statewide tests in reading, math, and science (beginning in 2007) in certain grades, with reports of such testing sent to parents. State standards in each assessed area must define levels corresponding to "basic" (considered to be below-standard), "proficient," and "advanced." Accordingly, under NCLB parents receive regular reports indicating whether their child is performing adequately based on evaluations external to
the child's classroom or school. (The NCLB Act and its accompanying federal regulations are very complex. The broad characterizations presented here have important nuances and exceptions that are too extensive to review in this chapter.)

An emphasis on external standards unequivocally is linked most closely with letter-grading practices that rely exclusively on evaluation of student products rather than process- or progress-oriented grading or the hodgepodge grading that is typically used (Cross & Frary, 1999). Descriptive standards-based grading, discussed earlier, is also consistent with the emphasis on external standards. Consider the disjuncture and confusion that would occur if a fifth-grade teacher awarded Bobby a B (signifying above average work) or even a C (average) in math based on his quiz scores, timely homework completion, and consistent good effort, only to have Bobby's parents receive a report from the end-of-year mandatory testing indicating only “basic” (i.e., less than adequate) achievement in mathematics. It's no surprise, then, that one large metropolitan area school district recently revised its grading policies for Grades 1–8 from a system previously based on attainment of “objectives assigned to the student” to one based on attainment of grade-level (Grades 1–5) or individual course (Grades 6–8) expectations, and based on achievement only (Montgomery County Public Schools, 2003).

Basing grades exclusively on attainment of uniform standards serves measurement functions well, but doing so may be seen as serving pedagogical functions (e.g., students' motivation, self-efficacy, etc.) less well (Buterman-Bos, Verloop, Terwel, & Wardekker, 2003). However, Buterman-Bos et al. found that secondary teachers in the Netherlands (where students must pass a national exam to graduate) employed "adjustment strategies," while still keeping an eye on eventual external-standards-based performance. Numerous adjustments in how students, parents, and teachers think about grades will need to occur as standards-based education moves ahead.

Effects of Different Grading Practices

What is known about the effects of grading practices is mostly a matter either of speculation or of inferences from theory that are not empirically tested. Although there is a vast literature on grading (over 4,000 published articles), most of it consists of essays and surveys of teachers' practices (McMillan & Workman, 1998). The empirical research that does exist tends to investigate both grading and assessment practices, making it difficult to draw conclusions about the effects of grading alone. However, comprehensive reviews of research (Crooks, 1988; McMillan & Workman, 1998) and individual studies (e.g., Ring & Reetz, 2000) lead to several conclusions. First, types of evaluations conducted (e.g., evaluations accenting broad, connecting principles vs. those emphasizing rote memorization of facts) affect both student study strategies and what students learn. Although this conclusion refers more to evaluation methods than to grading practices, it is well supported by research and strengthens the inference that grading practices may also affect student learning and motivation.

A second conclusion is that higher, but attainable, standards tend to lead to greater student effort and greater performance on achievement tests (Buterman-Bos et al., 2003; Crooks, 1988). The "but attainable" modifier is critical. Grades have some value as rewards, but no value as punishments: Low grades tend to cause students to withdraw from learning (Guskey & Bailey, 2001). Crooks (1988) summarizes this issue clearly, as follows: "In many teaching situations, [having high, but attainable standards] is not possible if all students are working simultaneously on the same tasks and trying to meet the same standards. Under such circumstances, some students will probably not be challenged, whereas others may find the standards unattainable" (p. 469).

By comparison to standards-based grading, which may have mixed effects, norm-referenced grading (i.e., grading systems that evaluate students relative to one another) is likely to be deleterious to far more students than it helps. Such evaluations decrease the learning and motivation of students who repeatedly score toward the bottom (Crooks, 1988) and are most motivating only for those toward the top of the distribution (Natriello, 1987).

A final conclusion is that some form of mastery grading, tied with a mastery learning approach to instruction, is well supported by the literature (Crooks, 1988; Guskey & Bailey, 2001). Mastery grading is a two-category (pass–fail) system, but mastery is typically defined at a very high level (corresponding to an A or B).

Mastery grading is inherently linked to the mastery learning approach. This approach to teaching uses both formative and summative evaluation, divides the curriculum into small units (e.g., about 2 weeks), and provides "corrective action" for students who did not reach mastery at the end of the unit. If unsuccessful on the first cycle of instruction and testing, students receive additional instruction and additional chances to demonstrate mastery. Mastery grading is the only grading method based specifically on a theory of student learning.
Mastery learning strategies have a strong record of improving both learning outcomes and students' attitudes at all educational levels (Guskey & Bailey, 2001), although the outcomes are much less clear in whole-school applications than in controlled studies (Haladyna, 1999). A disadvantage to the approach is that this integrated system requires extra time and work for teachers. This focus on integrated grading and instructional systems is consistent with the finding that although grading systems may affect students, instructional accommodations for low-achieving students likely have much greater impact than grading accommodations per se (Ring & Reetz, 2000).

Grading for Students With Disabilities

Students with disabilities often have difficulty learning and demonstrating mastery of information through assignments and traditional means of evaluation, such as tests and quizzes. This does not suggest that such students are incapable of learning the required content. Instead, these students often need some type of differentiated instruction, modifications, or accommodations to tests and assignments (e.g., reducing the number of required problems or giving extended time), which are typically delineated in their Individualized Education Program (IEP). Consequently, under traditional grading procedures students with disabilities may receive a grade for completing an assignment or test that is not the same as one completed by their peers without disabilities (Bradley & Calvin, 1998). To explore this complex topic, the following sections address (a) the need for adaptations, (b) effective adaptations currently used, (c) teacher collaboration and grading, (d) issues of fairness, and (e) application to low-achieving, non-special education students.

Need for adaptations. The increased focus on academic rigor and standards mandated by NCLB creates a sense of urgency to adhere stringently to the prescribed curriculum. As a result, it has rendered many teachers unwilling to differentiate their instruction or implement interventions that may compromise the integrity of the established standards. However, with the continuing increase in the number of students with disabilities who receive the majority of their instruction within general education environments, the individual needs of these students cannot be ignored. Those needs often require teachers to modify instruction as well as adapt their assessment and grading practices. Unfortunately, the demands associated with accountability for ensuring high student achievement, in conjunction with the requirement to meet the individual needs of students, pull teachers in different directions and force them to reconcile these requirements with their personal viewpoints about grading (McMillan & Nash, 2000).

Students with disabilities often are not successful in general education classes in which teachers do not adapt their grading practices. Students with disabilities in such classes either earn low grades or are at risk for failure (Munk & Bursuck, 2001). Moreover, students with disabilities in special education classes often receive higher grades than their peers with disabilities who are included in general education classes (Hendrickson & Gable, 1997). Because placing students with mild disabilities into self-contained special education classrooms is not consistent with the provisions and spirit of Individuals with Disabilities Education Act 2004, one possible method is for teachers to adapt their grading practices. Such adaptation is not a new concept: A survey of teachers revealed that half of general education teachers informally adapt their grading practices (Bursuck et al., 1996).

Munk and Bursuck (2001) contend that successful adaptation of grading practices depends in large part on the perceived purpose of grading, as discussed earlier in this chapter. Given the variety of perceived purposes, dissent among educators, parents, and students about adapted grading is not surprising. Additionally, the diverse purposes of grades are especially likely to cause conflict for students with disabilities. In the case of students with disabilities, hodgepodge grading may result in misleading grades and provide a false impression of the student's actual achievement level (McMillan et al., 2002).

Effective adaptations. Research has identified grading adaptations that teachers most commonly use. Specifically, Bursuck et al.'s (1996) survey of teachers revealed that the most commonly used grading adaptations acceptable to teachers included using weighted grades; using multiple grades, such as separate grades for effort and achievement; grading based on student progress and improvement; and grading based on mastery of IEP objectives. Conversely, the teachers reported the following as unacceptable: using adapted grading that evaluated students on mastery of less content than for general education students, altering the grading scale (e.g., lowering the percentage score required to earn an A), and promoting students to the next grade in the absence of achievement.
One interesting solution to the challenge of grading students with disabilities is to devise a Personalized Grading Plan, or PGP (Munk & Bursuck, 2001). Developing a PGP, which consists of multiple steps, could be completed within the IEP development and revision process to avoid additional meetings. The first step is for the student, parent, and teacher to jointly identify what purposes they believe the grade should serve. Second, these participants review school policy on grading and develop a menu of possible adaptations. Third, participants agree to the perceived purposes for a grade in order to select an appropriate adaptation. Fourth, the participants implement the PGP. Finally, the effects of the PGP on the individual student’s grades are evaluated, while soliciting the teacher’s perceptions of accuracy of those grades and discussing the student’s and parents’ satisfaction with the PGP.

Munk and Bursuck (2001) conducted a study on the implementation of a PGP with four eighth-grade students with learning disabilities. The results indicated that the students’ PGPs relied mostly on weighting the components differently (e.g., giving greater weight to homework versus tests) and slightly altering the criteria (percentage of points for a given grade), depending on the type of activity being evaluated. Additionally, the students’ grades mostly improved with the PGP or were higher than they would have been for that marking period had traditional grading been used. Overall, the students, parents, and teachers were generally satisfied with the PGPs because they achieved the intended purposes delineated at the outset. The results from this small study suggest a promising alternative to grading students with disabilities. However, even if it is incorporated within the IEP review process, it may not be feasible to implement PGPs for all students with disabilities, given the time investment required for all parties. Conversely, investing the time needed to clearly identify grading expectations and to reach an agreement among all participants on grading adaptations may, in fact, prevent misunderstandings and avert future conflicts. Additional suggestions for grading adaptations are listed in Table 1.

### Table 1  Suggested Grading Adaptations for Students With Disabilities

- Incorporate improvement and progress on IEP objectives into students' grades.
- Modify the weight of assignments, homework, tests, and quizzes.
- Have students create portfolios that incorporate examples of their completed work over time.
- Schedule frequent parent, teacher, and student conferences to discuss student progress and to address difficulties.
- Award grades based on student improvement.
- Use a point system for grading by which the student earns a specified number of points for assignments, homework, or tests and quizzes, which can be calculated daily, weekly, and throughout a marking period (see Hendrickson and Gable, 1997, for a comprehensive list and description of various grading options).
- Use points and percentages to grade modified assignments so that students can then be graded on work completed (e.g., if an assignment was modified to reduce the number of problems from 30 to 15).
- Avoid using different grading scales for students with disabilities (e.g., lower grading scale to allow them to earn an A).
- Avoid posting grades by identification number because it increases competition, and since students with disabilities often have the lowest scores, it may reinforce the notion that they are unable to achieve.
- Develop and use grading rubrics.
- Use a variety of grading approaches and methods.
- Avoid grading solely on effort.


benefits for both teachers and students. It offers the potential to enhance camaraderie among teachers, and discussion regarding instruction and grading can increase the consistency of expectations across teachers, which helps students who have multiple teachers.

The use of coteachers—teachers who work together within one classroom—represents another form of collaboration that is used across all grade levels. Coteaching by a special educator and a general educator in order to instruct and maintain students with disabilities in the general
education classroom now occurs frequently. Grading students with disabilities often presents a challenge for teachers, and it becomes even more complex with the involvement of more than one teacher, such as in a coteaching situation.

Grading traditionally is completed by individual teachers who employ their own philosophies. Most teachers have not been trained to grade collaboratively (Munk & Bursuck, 2001). As a result, many teachers have expressed confusion about their respective roles in terms of grading (Christiansen & Vogel, 1998). In an effort to address these difficulties, Christiansen and Vogel devised a decision model to be used by general and special educators who coteach and need to collaborate on grading for students with disabilities. This model is rooted in the problem-solving process and consists of four basic steps. First, the coteachers become aware of their school system grading policy, as well as state and federal policies. Second, each teacher identifies his or her basic philosophy of teaching and commonly used grading practices. Third, each teacher explicates his or her philosophy and grading practices to one another to become familiar with the other’s approach to teaching and grading. And fourth, after completing the first three steps, the coteachers cooperatively determine the appropriate grading practice for individual students.

**Fairness.** The issue of fairness has been noted throughout the research literature on grading practices, especially in reference to grading students with disabilities. Christiansen and Vogel (1998) delineated several reasons why many general education teachers are hesitant to alter their grading practices for students with disabilities. First, modifying grading for students with disabilities is seen as unfair to students without disabilities and potentially violates the school’s established grading policy. Teachers view allowing the practice of adapting grades only for students with disabilities as being particularly unfair to students without disabilities (Bursuck et al., 1996). Additionally, many general education teachers believe that the integrity of standards and established curriculum will be compromised if assignments and grading are modified for students with disabilities. This raises the question as to whether it is fair for students with disabilities, who have not mastered the same objectives or the same number of objectives, to earn a grade comparable to their peers without disabilities (Bradley & Calvin, 1998). The issue of comparable mastery becomes especially salient at the high school level, because states have specified graduation requirements that students must complete in order to earn a diploma. High school students are not only earning grades, but also are earning credits to fulfill the diploma requirements. The question thus becomes, is it fair for students with disabilities to receive credit for a class required for graduation in which they did not complete the same objectives as their peers without disabilities?

Ring and Reetz (2000) identified additional teacher beliefs about grading students with disabilities. In their research with teachers of students with learning disabilities within inclusive middle schools they cited four basic teacher beliefs. First, it is the individual student’s responsibility to achieve regardless of ability or disability. Second, accommodations are allowed based on student needs, but students should be held to grade level standards. Third, modifying assignments is acceptable. Fourth, grading can be based on effort and student functioning.

To obtain students’ viewpoints, Bursuck (1999) conducted a survey of high school students with and without disabilities in Grades 9–12 to determine what these students believed to be fair and acceptable grading adaptations. The results indicated that the responses differed based on level of achievement (high or low) and with respect to the presence or absence of a disability. Specifically, the majority of students without disabilities did not rate any grading adaptation as fair. Giving students passing grades in the absence of effort or corresponding achievement, using a modified grading scale, and giving a higher grade based on improvement were identified as the most unfair. However, if grading adaptations were to be used, those considered to be most fair (or least unfair) by all students included giving multiple grades (one to reflect effort and one for achievement) and passing students that put forth the effort. The most significant difference between high- and low-achieving students was that high-achieving students viewed weighting grades and calculating grade point average (GPA) based on difficulty of the class as fair (e.g., giving higher weights to grades in AP classes), whereas their low-achieving peers did not. Students with disabilities were much more inclined to perceive grading adaptations as acceptable for some students.

**Application to low-achieving students.** Much of the extant research, particularly that related to students with learning disabilities, has not identified educationally relevant differences between low-achieving students and school-identified students with learning
disabilities (Stuebing et al., 2002). In fact, many low-achieving students have similar needs and experience similar educational difficulties as their peers identified with learning disabilities. Teachers of low-achieving students may find it beneficial to implement grading practices similar to those they employ with students with disabilities. However, a critical challenge will likely arise: Entitlements under IDEA 2004 do not extend to low-achieving students who are not formally identified as having an educational disability. Accordingly, issues of the appropriateness and fairness of using adapted grading with such students will loom large in the absence of mandated modifications and accommodations.

RECOMMENDED GRADING PRACTICES

Given the functions of grading, both historically and currently, the real-world competing pressures on teachers, the mandates of NCLB, and advice from measurement experts, several recommendations can be made regarding grading.

Articulate a clear policy on grading and disseminate it to parents and students. School systems should inform teachers, parents, and students about policies pertaining to grading so that consumers can correctly interpret the meaning of the grades. Articulate the bases allowed for grading. Say whether adaptations are allowed and, if so, for whom or under what circumstances. To the degree that teachers have some latitude in grading procedures, they should make clear their expectations and the basis on which grades will be assigned. Of course, this must be done in a developmentally appropriate fashion; information given to secondary school students will be much different than that given to those in early elementary school.

Integrate assessment and grading practices. Although grades represent an evaluation of student achievement, they should be based solidly on valid assessments of that achievement and reflect clearly articulated instructional objectives. Grading and assessment practices need to be aligned. For example, a grading system that reports on specifics such as "student can write about material that he or she has read" must have some valid measure of assessment that links reading and writing beyond the teacher's mere global impression of student achievement.

Ensure that grades that refer to academic content predominantly, or exclusively, reflect current achievement. If grades are to communicate clearly, a grade or other performance indicator in math should reflect the student's math achievement, not interest or effort or improvement, unless such intended nonachievement meanings are made explicit on the report card or on the student's work. Grades that indicate achievement should weight comprehensive sources of information (e.g., tests vs. quizzes) and the most recent evidence most heavily.

Use multiple grades or indicators rather than large composite grades. Multiple indicators take two forms: (a) multiple indicators of achievement within an area, and (b) separate indicators for achievement and nonachievement outcomes. For example, rather than give a student one letter grade for math, a teacher might grade math calculation and math applications separately. This scheme increases the specificity of information provided to students and parents and decreases the inaccuracies that may result from averaging scores across disparate learning objectives. Similarly, report cards might include letter grades strictly for achievement and give a separate set of indicators for outcomes such as effort, study skills, attendance, and so forth.

Use a criterion-referenced system at least through elementary school. Systems that evaluate student achievement against externally set standards will most clearly communicate information about student accomplishment. There is little evidence supporting the value or desirability of grading that focuses on normative evaluations such as superior, above average, or average (i.e., traditional interpretations of ABCDF grading), at least until there is a need for the school system to provide interindividual comparisons for competitive selection processes, such as college admissions.

Develop a system for grading students with disabilities that is consistent with accommodations provided and/or with differential expectations. To the extent that the educational program for a student with disabilities is altered from the program for all other students, the grading system for that student should reflect such changes.
SUMMARY

Grading and reporting is an essential feature of American education. How this function is performed has significant implications for the education and mental health of K–12 students. Educators, and the psychologists who work with them, need a better understanding of grading and reporting in order to improve this aspect of the educational process. Challenges abound and include (a) improving the alignment of assessment and grading practices with established measurement and evaluation principles; (b) developing reporting systems that, in the modern age of standards and high-stakes testing, accurately reflect the student’s achievement status relative to mandated, external-to-the-classroom standards but do not drastically undermine the motivation of lower achieving students; and (c) grading students with disabilities fairly and accurately in the context of programs that specifically include accommodated instruction and individualized goals.

Grading and reporting almost certainly will continue to be one of the most controversial and passionately debated issues in education. Combining knowledge of the larger context in which grading and reporting is embedded with an understanding of stakeholders’ views and sound measurement and evaluation principles should allow school psychologists and other educational consultants to navigate thoughtfully through this difficult terrain.

REFERENCES


Children’s Needs III


