

Principal's Research Review

Supporting the Principal's Data-Informed Decisions

ISSN 1558-5948

VOL. 4, ISSUE 1

JANUARY 2009

Reforming Grading Practices in Secondary Schools

By Ken O'Connor

The primary purpose for grading...should be to communicate with students and parents about their achievement of learning goals.

Susan M. Brookhart, *Grading*

Although grades are not essential for learning, they dominate the culture of high schools. Every state has standards, but high schools are only reluctantly becoming standards-based for curriculum, instruction, and assessment and few are standards-based for grading and reporting. Only rarely has grading been part of preservice or inservice training for teachers. Grading can be idiosyncratic, private, and based largely on how a teacher experienced grading as a student or a young professional. As a result, "neither the weight of scholarship nor common sense seems to have influenced grading policies in many schools. Practices vary greatly among teachers in the same school—and even worse, the practices best supported by research are rarely in evidence" (Reeves, 2008, p. 85). For this review, a grade is a symbol (letter or number) on a report card that summarizes student achievement. A mark or score is the symbol (letter or number) given to any student test or performance that provides evidence of student achievement.

Understanding Grading Practices

What grading practices are in place in secondary schools and what are the most appropriate practices? The most recent comprehensive survey on high school grading policies was published in 1998 by the College Board. It found that:

a large majority of schools use a traditional grading system of A–F or numeric grades (91 per cent), use the same grading system for all academic courses (92.2 per cent), report GPA (90.1 per cent), and calculate a high school class rank (81.3 per cent).... Approximately 8 per cent of schools report using a nontraditional grading system and only 1 per cent of schools do not assign grades." (p. 2)

There has been little change since then except that fewer schools are reporting class rank.

Stiggins, Frisbie, and Griswold (1989) identified 19 recommended grading practices and also gathered information on the grading practices of a number of secondary teachers. They found no discrepancy on 8 practices but did find discrepancy on 11 others, including

When teachers emphasize assessment for learning, student achievement improves, student ownership of learning improves, and student engagement increases.

Absolum (2006)

recommendations that achievement should be the sole ingredient in grades, that attitudes should not be included as a grading variable, and that consistent policies should be followed. (See figure 1.) Stiggins et al. (1989) suggested three possible reasons for the discrepancies: that best practice is a matter of opinion, that measurement specialists fail to take into account the practical realities of the classroom, and that teachers are unaware of the recommendations. A consensus has since emerged as to what is best practice, and a number of researchers and practitioners have recommended guidelines for grading (e.g., Guskey & Bailey, 2001; Tomlinson & McTighe, 2006; Marzano, 2000, 2006; Stiggins, Arter, Chappius, and Chappius, 2004; Wormeli, 2006; Cooper, 2007; and O'Connor, 2002, 2007). There are differences in order, emphasis, and words in the guidelines, but these educators have many years of teaching experience and make similar recommendations.

Some researchers, such as Kohn (1994), called for abandoning grades completely. He suggested that grades lead to less successful learning, less interest in

learning, and less willingness to engage in challenging tasks. However, Clymer and Wiliam (2006/07) disagreed. They stated that if teachers do not provide some indication of students' achievement, school systems are likely to rely on timed written examinations. They also found that appropriately designed grading systems can help identify where students are in their understanding and what they need to improve.

It is essential to be clear about the primary purpose of grades, which is to communicate students' achievement of learning goals. As Brookhart (2004) noted, grades have a secondary purpose that includes providing teachers with information for instructional planning and providing teachers, administrators, parents, and students with information for placement of students. She also noted that the main difficulty driving grading issues is that grades serve a variety of conflicting purposes. Bailey and McTighe (1996) agreed that the primary purpose of grades is to communicate student achievement to students, parents, school administrators, postsecondary institutions, and employers.

Figure 1.

Recommended Grading Practices

Discrepancy

- Achievement as only characteristic in grades
- Ability not included in grades
- Motivation and effort included in grades
- All daily assignments included in grades
- Amount of grading data gathered
- Quality of grading data
- Consistent policies followed
- Methods of aggregating components
- Fixed percentages as cutoff scores
- Total point accumulation for cutoff scores
- Deciding on borderline cases

No Discrepancy

- Communicating grading methods to students
- Attitude not in grades
- Interest not in grades
- Personality not in grades
- Written tests in grades
- Oral questioning during instruction not in grades
- Performance assessments in grades
- Normal distribution not used

Stiggins, Frisbie, and Griswold (1989, pp. 7-8)

Major Issues of Grading

Grading Variables

Most teachers have combined achievement with behavior to varying extents in determining grades because they believe it demonstrates what they value and will motivate students to exhibit those behaviors.

McMillan (2001) noted that “the findings from this study, along with other results from other studies, show that this practice is still pervasive” (p. 30).

Gathercoal (2004) noted that “due to the excessive entanglement between *achievement* and *behavior*, achievement grades are often misinterpreted” (p. 153).

Tony Winger, a high school teacher in Littleton, CO, (2005) said:

I recall telling my students, “Work hard and your grade will be fine.” Although I did not realize it, the message to my students was clear: My unconscious message to my students was one of compliance.... Some students received good grades and learned little, others learned much and failed (p. 62).

Reeves (2008) concluded, “When schools improve grading policies—for example, by disconnecting grades from behavior—student achievement increases and behavior improves dramatically” (p. 90).

According to Guskey (2009), the criteria for grades can be grouped into three broad categories: *product, process, and progress.* Product criteria is concerned with what students know and are able to do at a particular point in time. Process criteria looks at how students got there, which includes work habits, homework, punctuality of assignments, and class participation. Progress criteria focus on what students gained from their learning experiences.

Guskey found that most teachers base their grading procedures on a combination of all three types of criteria. He also noted that most researchers and measurement specialists recommend the exclusive use of product criteria in determining students’ grades. Given the concerns that lead to these differences, Guskey recommended a solution used by “increasing numbers of students and schools...to report

separate grades or marks for students on each set of criteria” (p. 20). This requires the use of expanded report cards that provide a grade for achievement only (product), specific information on learning skills (process), and sufficient space for teachers to be able to comment on growth (progress).

Guskey and Bailey (2001) identified typical sources of grading and reporting evidence: exams or compositions, quizzes, reports or projects, student portfolios, oral presentations, homework, class participation, laboratory projects, student notebooks or journals, and punctuality of assignments. Other grading sources include effort, attendance, behavior, and progress made.

Purpose of Assessment

Welsh and D’Agostino (2009) stated, “It is critical that teachers understand differing roles of assessment and that they effectively select assessments to meet a particular need” (p. 102). They also noted that “while all assessment results should be reviewed by teachers, it is not appropriate to incorporate all forms of assessment in grading” (p. 102). There are three types of assessment: diagnostic, formative, and summative assessment. Diagnostic assessment, or pre-assessment, occurs before instruction to guide teachers in their planning of instruction based on what the assessments reveal about their students’ existing knowledge and skills. Formative assessment, or assessment *for* learning, occurs during the learning process and provides information to both teachers and students that enables them to make adjustments to increase learning. Summative assessment, or assessment *of* learning, occurs after instruction to find out what students know, understand and can do at one point in time. Welsh and D’Agostino emphasized that grades should be based on summative assessments only and that diagnostic and formative assessment should not contribute directly to grades. Students should not be evaluated on content they have not studied or penalized while attempting to learn new content.

A wealth of research supports this point. A definitive study by Black and Wiliam (1998) reviewed 580 articles or chapters on assessment from 1989–98 and concluded that that there is overwhelming evidence

that improving formative assessment raises standards, that formative assessment needs improvement, and that the evidence provides direction about how to improve formative assessment. Improving formative assessment helps all students, but the largest learning gains occur for low achievers. Among the problems were that marks and grades were overemphasized with little advice to students about how to improve. Feedback to students often seemed “to serve social and managerial functions, often at the expense of the learning function” (p. 142). Formative assessment can be improved by providing opportunities for self-assessment, which Black and Wiliam argued “is in fact *an essential component of formative assessment*” (p. 143). They further stated that “When anyone is trying to learn, feedback about the effort has three elements: recognition of the *desired goal*, evidence about *present position*, and some understanding of a way to *close the gap* between the two” (p. 143). Black and Wiliam stated that “feedback has been shown to improve learning when it gives each pupil specific guidance on strengths and weaknesses, preferably without any overall marks” (p. 144).

The last point needs to be emphasized in professional development and in school and district grading policies: to be effective, formative assessment must be no mark, comment only. Black, Harrison, Lee, Marshall, and Wiliam (2003) reported that “initial fears about how students might react to not receiving marks turned out to be unjustified” (p. 45), and that students saw the relationship between effort and improved learning.

The key to success is feedback. “The quality of the feedback rather than its existence or absence is the central point” (Atkin, Black, & Coffey, 2001, p. 15). Davies (2007) listed the characteristics of effective feedback. Descriptive feedback:

- Comes during as well as after learning
- Is easily understood and relates directly to the learning
- Is specific, so performance can improve
- Involves choice on the part of the learner as to the type of feedback and how to receive it
- Is part of an ongoing conversation about the learning

- Is in comparison to models, exemplars, samples, or descriptions (and)
- Is about the performance or the work—not the person. (p. 17)

When teachers emphasize assessment for learning, student achievement improves, student behavior improves, student ownership of learning improves, and student engagement increases. (Absolum, 2006).

Clear Learning Goals

Although they have different labels (standards, learning results, expectations, and outcomes), every state has standards that are determined at the state level. These standards are published and all teachers, parents, and students, should be familiar with them. This is essential because the research shows that “it is very difficult for students to achieve a learning goal unless they understand that goal and can assess what they need to do to reach it” (Black et al., 2003, p. 49).

McMillan stated, “The promise of standards-based grading is that both teachers and students will have a clearer conception of what needs to be learned and of what constitutes successful performance” (2009, p. 107). He recognized what is often not acknowledged—standards have two components—the “what” (content standards) and the “how well” (performance standards) and both must be clearly defined.

“Grades typically carry little meaning because they reduce a great deal of information to a single letter” (Atkin et al., 2001, p. 64). As Trumbull and Farr noted in standards-based systems, assessments often “employ scoring systems that rate students on different aspects of performance. If writing is evaluated according to sub-domains like ‘content/ideas,’ ‘cohesion/structure,’ and ‘mechanics,’ then to reduce scores on these three scales to a single grade is to obscure important performance differences” (Trumbull & Farr, p. 29).

In standards-based systems, teachers should move from an assessment methods-based system to a standards-based system where the categories in the gradebook are not tests, projects, and assignments but, for example, a classification similar to what Winger (2005) used in his *Introduction to Sociology*

class. "I grouped essential academic expectations into four components: conceptual understanding, application, analysis and evaluation, and formal writing" (Winger, p. 63). With multiple scores on tests and assignments, a picture can be built of each student's achievement in each category and summarized at the end of the grading period. In most secondary schools, this information will then be reduced to a single grade, but the most valuable information is provided by the profile and it is essential that this profile be provided to students and parents on standards-based report cards. Benson (2008) stated, "In standards-based schools, grades are replaced with, or augmented by, achievement reports that indicate *levels of performance on essential benchmarks* [italics added]" (p. 35).

If secondary schools are going to be truly standards-based, they need to determine what constitutes passing or receiving a credit. It is highly likely that a student will be proficient on some standards and far from proficient on other standards—if this is the case, should they get a credit? A number of high schools are addressing this by defining credits as proficiency on most standards and partial proficiency on the remaining standards. Foxcroft Academy in Maine requires students to be at least proficient on most standards and partially proficient on the others.

Grading and assessment are basically about "how good is it" or "is it good enough" so we must also have clear performance standards. Traditionally, points have provided performance standards in secondary schools with grades then being determined on a percentage scale. Madgic (1988) identified a number of problems with this approach: "misplaced emphasis," "illusion of objectivity," "reduction of teacher judgment and responsibility," "cumulative point totals and cumulative errors," and "fallacies of 'standard' percentage categories" (pp. 30–31). Madgic (1988) stated that the latter is:

the most glaring deficiency of a "standard" percentage approach (90–100 = A, etc.) is the presumption that a certain percentage represents a valid rating of a performance

level, and that a teacher can decide on these percentage categories in advance. This presumption is certainly not true unless it [assessment/test] has been evaluated...so that its results...represent a...valid indicator of student performance levels. (p. 31)

Clearly, the percentage system and a standards-based system are incompatible. In a true standards-based system, performance standards are based on proficiency and it is then necessary to decide how many levels there will be above and below proficiency. There is no right number of levels—the Advanced Placement program uses five levels while the International Baccalaureate uses seven levels—but it is probably closer to two (proficient/not proficient) than 101 (the percentage system).

Motivation

In traditional grading practices, grades have often functioned as rewards and punishments to motivate students to achieve and behave appropriately. Penalties for late work, zeros for missing assignments and academic dishonesty, and inclusion of behaviors as part of grades were employed to promote student accountability and responsibility—all based on extrinsic motivation. Currently most secondary schools aim to develop students to be self-directed, independent, lifelong learners. This will not happen if educators rely on extrinsic motivation—they must instead establish approaches that maximize intrinsic motivation. As Manitoba Education (2006) stated,

According to current cognitive research, people are motivated to learn by success and competence. When students feel ownership and have choice in their learning, they are more likely to invest time and energy in it. Assessment can be a motivator, not through reward and punishment, but by stimulating students' intrinsic interest. Assessment can enhance student motivation by:

- Emphasizing progress and achievement rather than failure

- Providing feedback to move learning forward
- Reinforcing the idea that students have control over, and responsibility for, their own learning
- Building confidence in students so they can and need to take risks
- Being relevant, and appealing to students' imaginations
- Providing the scaffolding that students need to genuinely succeed. (p. 7)

Stiggins et al. (2004) noted that intrinsic motivation can be diminished by coercion, intimidation, rewards or punishments linked to grades, infrequent or vague feedback, limitation of personal control, and responsibility without authority.

Dweck (2007) said, "It matters greatly what students believe about their intelligence" (p. 6). She distinguished between students with a "fixed mindset" who believe that intelligence is innate and unchangeable and those with a growth mindset who believe that their achievement can improve through effort and learning. According to Dweck, recent studies show "that teaching students a growth mindset results in increased motivation, better grades, and higher achievement test results" (p. 10).

Guskey (2009) noted, "no studies support the use of low grades as punishment. Instead of prompting greater effort, low grades more often cause students to withdraw from learning" (p. 14). Motivation is enhanced when students are provided accurate information about achievement, have clear learning goals, and study in an environment that supports learning by not including diagnostic and formative assessment in grades and by being positive and supportive, not negative or punitive.

Summary

Grading in secondary schools in standards-based systems is complex, partly because of the hold of traditional grading practices. To make grades accurate, meaningful, consistent, and supportive of learning:

- Grades must be about achievement with behaviors reported separately
- Grades must be determined primarily from summative assessments

- Formative assessment should be no mark, comment only and provide clear, specific, descriptive feedback
- Learning goals—both the *what* and the *how well*—must be clear for teachers, students, and parents
- Emphasis must be placed on intrinsic motivation.

For secondary schools to achieve their mission of proficiency for all and developing students into self-directed, independent, lifelong learners, grades must be seen as communication tools and not as motivators. This means that grades must be about achievement only with behaviors reported separately, and the processes that lead to their determination must support the learning process so that students understand that school is about learning and not just accumulating points. PRR

References

- Absolum, M. (2006). *Clarity in the classroom: using formative assessment, building learning-focused relationships*. Auckland, NZ: Hodder Education.
- Atkin, J. M., Black, P., & Coffey, J. (Eds.). (2001). *Classroom assessment and the national science education standards*. Washington, DC: National Academy Press.
- Bailey, J., & McTighe, J. (1996). Reporting achievement at the secondary school level: What and how? In T. R. Guskey, (Ed.), *Communicating student learning: ASCD Yearbook 1996* (pp. 119–140). Alexandria, VA: ASCD.
- Benson, D. (2008). *The standards-based teaching/learning cycle*. The Colorado Coalition for Standards-Based Education.
- Black, P. J., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2003). *Assessment for learning: Putting it into practice*. Maidenhead, England: Open University Press.
- Black, P. J., & Wiliam, D. (1998). Inside the black box. *Phi Delta Kappan*, October, 80(2), 139–148.
- Brookhart, S. M. (2004). *Grading*. Upper Saddle River, NJ: Pearson Merrill Prentice Hall.
- Clymer, J. B., & Wiliam, D. (2006/2007). Improving the way we grade science. *Educational Leadership*, 64(4), 36–42.

- The College Board. (1998, May). High school grading policies. *Research Notes*, RN-04, New York: Author.
- Cooper, D. (2007). *Talk about assessment*. Toronto, ON, Canada: Nelson.
- Davies, A. (2007). *Making classroom assessment work* (2nd ed.). Courtenay, BC: Connections Publishing.
- Dweck, C. S. (2007). Boosting achievement with messages that motivate. *Education Canada*, 47(2), 6–10.
- Gathercoal, F. (2004). *Judicious discipline*. (6th ed.). San Francisco: Caddo Gap Press.
- Guskey, T. R. (2009). Grading policies that work against standards...and how to fix them. In T. R. Guskey (Ed.), *Practical Solutions for Serious Problems in Standards-Based Grading* (pp. 9–26). Thousand Oaks, CA: Corwin.
- Guskey, T. R., & Bailey, J. (2001). *Developing grading and reporting systems for student learning*. Thousand Oaks, CA: Corwin Press.
- Kohn, A. (1994). Grading: The issue is not how but why. *Educational Leadership*, 52(2), 38–41.
- Madgic, R. F. (1988). The point system of grading: A critical appraisal. *NASSP Bulletin*, 72(507), 29–34.
- Manitoba Education, Citizenship and Youth. (2006). *Rethinking classroom assessment: Assessment for learning, assessment as learning, assessment of learning*. Winnipeg, MB: Author.
- Marzano, R. J. (2000). *Transforming classroom grading*. Alexandria, VA: ASCD.
- Marzano, R. J. (2006). *Classroom assessment and grading that work*. Alexandria, VA: ASCD.
- McMillan, J. H. (2001). Secondary teachers classroom assessment and grading practices. *Educational Measurement: Issues and Practices*, 20(1), 20–32.
- McMillan, J. H. (2009). Synthesis of issues and implications for practice. In T. R. Guskey (Ed.), *Practical Solutions for Serious Problems in Standards-Based Grading* (pp. 105–120). Thousand Oaks, CA: Corwin.
- O'Connor, K. (2002). *How to grade for learning: Linking grades to standards* (2nd ed.). Thousand Oaks, CA: Corwin.
- O'Connor, K. (2007). *A repair kit for grading: 15 fixes for broken grades*. Portland, OR: ETS.
- Reeves, D. B. (2008). Effective grading practices. *Educational Leadership*, 65(5), 85–87.
- Reeves, D. B. (2008). Improving student attendance. *Educational Leadership*, 65(8), 90–91.
- Stiggins, R. J., Arter, J., Chappuis, J., & Chappuis, S. (2004). *Classroom assessment for student learning: Doing it right, using it well*. Portland, OR: ETS.
- Stiggins, R. J., Frisbie, D. A., & Griswold, P. A. (1989). Inside high school grading practices: Building a research agenda. *Educational Measurement: Issues and Practice*, 8(2), 5–14.
- Tomlinson, C. A., & McTighe, J. (2006). *Integrating differentiated instruction and understanding by design*. Alexandria, VA: ASCD.
- Trumbull, E., & Farr, B. (2000). *Grading and reporting in an age of standards*. Norwood, MA: Christopher-Gordon.
- Welsh, M. E., & D'Agostino, J. (2009). Fostering consistency between standards-based grades and large-scale assessment results. In T. R. Guskey (Ed.), *Practical solutions for Serious Problems in Standards-Based Grading* (pp. 75–104). Thousand Oaks, CA: Corwin.
- Winger, T. (2005). Grading to communicate. *Educational Leadership*, 63(3), 61–65.
- Wormeli, R. (2006). *Fair isn't always equal: Assessing and grading in the differentiated classroom*. Portland, ME: Stenhouse/NMSA.

About the Author

Ken O'Connor is an independent consultant on assessment, grading, and reporting, who has been a presenter and facilitator in 40 states, 8 provinces in Canada, and in 11 countries. He is the author of *How to Grade for Learning: Linking Grades to Standards* (2009, Corwin) and *A Repair Kit for Grading: 15 Fixes for Broken Grades* (2007, ETS/ATI).