

Grading Policies that Work Against Standards...and How To Fix Them

Thomas R. Guskey

Thomas R. Guskey is a professor in the Educational Policy Studies and Evaluation department, College of Education, University of Kentucky, Lexington. His course offerings center on quantitative analysis, research and evaluation design, educational change, and instructional quality. Correspondence concerning this article may be sent to guskey@pop.uky.edu. Copyright 2000 National Association of Secondary School Principals

An important element of a successful standards-based reform initiative includes grading and reporting that refers to specific learning criteria rather than normative criteria. Four grading policies that impose barriers to reform are described. Specific strategies to correct these policies are offered.

Most educators welcome the current reform efforts that focus on standards. By providing consensus about what's important for students to learn and what skills they should acquire, standards give direction to modern reform initiatives. In particular, they bring much needed focus to curriculum development work and provide the impetus for fashioning new forms of student assessment.

If the true benefits of standards are to be realized, however, educational leaders must view their reform initiatives systemically. This means that in addition to essential curriculum and assessment issues, leaders also must consider organizational factors that exert potentially strong influence on implementation. Policies and organizational procedures at the district, school, and classroom levels can profoundly impact reform initiatives and significantly affect results. Research indicates (see Lieberman 1995) the most carefully articulated curriculum and best-aligned assessments will make little difference if school policies stand in the way of implementation.

Described in this article are four school policies that impose procedural barriers to the implementation of standards-based reforms. Also described are specific strategies for correcting them. Each of these policies relates to grading and reporting practices; that is, how students' learning progress is summarized and communicated to parents, students, and others. Despite their importance, grading and reporting are seldom included in discussions of curriculum or assessment reform. Nevertheless, their powerful influence can prevent even modest success in any standards-based reform initiative.

Policy 1: Grading "On the Curve"

In a standards-based system, grading and reporting must be done in reference to specific learning criteria, rather than in reference to normative criteria or "on the curve." In other words, students must be graded in terms of what they have learned and are able to do, not in terms of their relative standing among classmates. The principal advantage of using the normal distribution curve as a basis for assigning grades is that it ensures consistent grade distributions from one teacher to the next. Consequently, every teacher's classes have the same percentage of As, Bs, Cs, etc. But the consequences of this practice are overwhelmingly negative. Research indicates that it is detrimental to the relationships among students and to the relationships between teachers and students (Krumboltz and Yeh 1996).

Grading on the curve makes learning a highly competitive activity in which students compete against one another for the few scarce rewards (high grades) distributed by the teacher. Under these conditions, students readily see that helping others become successful threatens their own chances for success (Gray 1993; R. T. Johnson, Johnson, and Tauer 1979; D. W. Johnson, Skon, and Johnson 1980). High grades are attained not through excellence in performance, but simply by doing better than one's classmates. As a result, learning becomes a game of winners and losers, and because the number of rewards is kept arbitrarily small, most students are forced to be losers (Haladyna 1999; D. W. Johnson and Johnson 1989).

Most students, as well as most adults, can relate horror stories based on their experiences in classes where they were graded on the curve. Many recall the anger they felt toward the high scoring student in their class who “inflated the curve” and, in their minds, caused other class members to receive a lower grade. Some remember being the object of their classmates’ anger because they were that high scoring student. Stories also abound of students hiding books in the library so that their classmates could not use them or removing equipment needed in projects or experiments in order to enhance their chances for a high grade. Furthermore, grading on the curve denies students the opportunity to work together and to help each other attain valuable, shared learning goals.

Perhaps most important, grading on the curve communicates nothing about what students have learned or are able to do. Rather, it tells only a student’s relative standing among classmates, based on what are often ill-defined criteria. Students who receive the high grades might actually have performed very poorly in terms of the established learning standards, but simply less poorly than their classmates. Differences between grades, therefore, are difficult to interpret at best, and meaningless at worst (Bracey 1994).

If the purpose of grading is to reflect what students have learned and are able to do, then grading on the curve falls far short. As Bloom, Madaus, and Hastings (1981) so succinctly put it:

There is nothing sacred about the normal curve. It is the distribution most appropriate to chance and random activity. Education is a purposeful activity, and we seek to have students learn what we have to teach. If we are effective in our instruction, the distribution of achievement should be very different from the normal curve. In fact, we may even insist that our educational efforts have been unsuccessful to the extent that the distribution of achievement approximates the normal distribution (52-53).

Other unintended but equally adverse consequences for students can result from grading on the curve. A study by Wood (1994), for example, found the percentage of students receiving particular grades in an urban high school remained virtually the same from the sophomore through senior years. At first glance this appears to show that teachers throughout the school were remarkably consistent in their grading. However, Wood also found that each year there were fewer students in the school. Because students who leave are generally those with the lowest grades, this consistency in grade percentages means that as one group of unsuccessful students drops out, it is replaced by a succession of newly created low grade students who were formerly successful. In other words, additional students are at risk of failing each year. Some students who got Cs as sophomores will get Ds as juniors, and so on

Furthermore, modern research has shown that the seemingly direct relationship between aptitude or intelligence and school achievement depends on instructional conditions, not a normal distribution curve (Engel 1991). When the instructional quality is high and well matched to students’ learning needs, the magnitude of this relationship diminishes drastically and approaches zero (Bloom 1976; Bloom, Madaus, and Hastings 1981). Moreover, the fairness and equity of grading “on the curve” is a myth.

Remedy

In any educational setting where the central purpose is to have students learn, grading and reporting should always be done in reference to specific learning criteria, rather than in reference to normative criteria. Because normative criteria or grading on the curve tells nothing about what students have learned or are able to do, they provide an inadequate description of student learning. In addition, they promote unhealthy competition, destroy perseverance and other motivational traits, and are generally unfair to students (Haladyna 1999). At all levels of education, therefore, teachers should identify what they want their students to learn, what evidence they will use to verify that learning, and what criteria they will use to judge that evidence. In other words, teachers should clarify their standards and their grading criteria on the basis of those standards. Grades based on specified learning criteria and standards have direct meaning

Policy 2: Selecting Valedictorians

Although many educators today understand the negative consequences of grading on the curve and have abandoned the practice, most fail to recognize other common school policies that yield similar negative consequences. One of the most prevalent is the way in which schools select class valedictorians. There is nothing wrong, of course, with recognizing excellence in academic performance. But in selecting the class valedictorian, most schools operate under the traditional premise that there can be only one. This commonly results in severe and sometimes bitter competition among high achieving students to be that one. Early in their high school careers top students figure out the selection procedures and then, often with the help of their parents, find ingenious ways to improve their standing in comparison to classmates. Again, to gain that honor a student must not simply excel; he or she must outdo the other students in the class. And sometimes the difference among these top students is as little as one-thousandth of a decimal point in a weighted grade-point average.

Remedy

An increasing number of high schools have resolved this problem simply by moving away from the policy of having just one valedictorian and, instead, naming multiple valedictorians. This is similar to what colleges and universities do in naming graduates cum laude, magna cum laude, and summa cum laude. West Springfield High School in Fairfax County, Va., for example, typically graduates 15 to 25 valedictorians each year (Smith 1999). Every one of these students has an exemplary academic record that includes earning the highest grade possible in numerous honors and Advanced Placement classes. Instead of trying to distinguish among these exceptional students, the faculty at West Springfield High School decided that all should be named valedictorians. In other words, rather than creating additional, arbitrary criteria in order to discriminate among these high-achieving students (considering, for example, their academic record from middle school or even elementary school), they decided to recognize the excellent achievement and performance of the entire group. And because the faculty at West Springfield High School believes their purpose as teachers is not to select talent but rather to develop it, they take great pride in these results. All of the valedictorians are named at the graduation ceremony, and one student, selected by his or her fellow valedictorians, makes a major presentation.

Some might object to a policy that allows multiple valedictorians, arguing that colleges and universities demand such selection and often grant special scholarships to students who attain that singular distinction. But current evidence indicates this is not the case. In processing admission applications and making decisions about scholarships, college and universities are far more interested in the rigor of the curriculum students have experienced (Bracey 1999). In fact, an index composed of the number of Advanced Placement courses taken, the highest level of math studied, and total number of courses completed has been shown to be a much stronger predictor of college success than standardized test scores, grade point average, or class rank (Adelman 1999). The rigor of the academic program experienced by the valedictorians from West Springfield High School has helped them gain admission and win scholarships to many of the most selective colleges and universities in the nation.

The process by which class valedictorians are selected is another example of a policy that continues not because educators have thought about it deeply, but simply because they have “always done it that way.” It is also a policy that hinders the implementation of standards-based reforms. Better understanding of the consequences of such a policy allows education to implement improved and more appropriate policies that benefit students and teachers alike.

Recognizing excellence in academic performance is a vital aspect in any learning community. However, such recognition need not be based on arbitrary standards and deleterious competition. Instead, it can and should be based on clear models of excellence that exemplify our highest standards and goals for students and for ourselves. And if many students meet these high standards of excellence, all the better.

Policy 3: Using Grades as a Form of Punishment

Although educators would undoubtedly prefer that motivation to learn be entirely intrinsic, grades and other reporting methods are important factors in determining how much effort students put forth (Cameron and Pierce 1994, 1996; Chastain 1990; Ebel 1979). Studies show that most students view high grades as positive recognition of their success, and some work hard to avoid the consequences of low grades (Feldmesser 1971).

At the same time, no studies support the use of low grades or marks as punishments. Instead of prompting greater effort, low grades more often cause students to withdraw from learning. To protect their self-images, many students regard the low grade as irrelevant and meaningless. Other students may blame themselves for the low grade, but they may feel helpless to make any improvement (Selby and Murphy 1992).

Sadly, some teachers consider grades or reporting forms as their “weapon of last resort.” In their view, students who do not comply with their requests must suffer the consequences of the greatest punishment a teacher can bestow: a failing grade. Such practices have no educational value and, in the long run, adversely affect students, teachers, and the relationship they share.

Remedy

Rather than attempting to punish students with a low grade or mark in the hope it will prompt greater effort in the future, teachers can better motivate students by considering their work as incomplete and then requiring additional effort. Recognizing this, some schools have initiated grading policies that eliminate the use of failing grades altogether. Teachers at Beachwood Middle School in Beachwood, Ohio, for example, record students’ grades as A, B, C, or I (Incomplete). Students who receive an I grade are required to do additional work in order to bring their performance up to an acceptable level. This policy is based on the belief that students perform at a failure level or submit failing work in large part because teachers accept it. If teachers no longer accept substandard work, Beachwood educators reason, then students will not submit it and, with appropriate support, will continue to work until their performance is satisfactory. (Bernetich 1998). Beachwood Middle School teachers strongly believe that giving a failing grade to students who have not performed well, despite their ability to do so, offers these students an easy way out of schoolwork. By contrast, if teachers insist that all assignments designed to demonstrate learning be completed and done well, then students will choose to do their work in a timely fashion and at a satisfactory level of quality. The guiding maxim of the teachers at Beachwood Middle School is “If it’s not done well, then it’s not done!”

Implementing grading policies such as this naturally requires additional funding for the necessary support mechanisms. Students who receive an I grade at Beachwood, for example, are required to attend after-school sessions or special Saturday school programs staffed by teachers, volunteer parents, and older students. Those who are unable or unwilling to do the make-up work during the school year must attend required summer school sessions designed to help them bring their performance up to an acceptable level (Kuehner 1998). Although these support mechanisms demand commitment and additional funding, schools implementing such programs generally find them to be highly successful (Bernetich 1998). Many also discover that they actually save money in the long run. Because this regular and ongoing support helps students remedy their learning difficulties before they become major problems, less time and fewer resources need to be spent in major remediation efforts later on.

At all levels of education, we need to think seriously about the use of failing grades. Although honesty must prevail in assessment and evaluation of student learning, we also must consider the negative consequences of assigning failing grades to students’ work or level of performance (see Roderick and Camburn 1999). Especially in the early years of school, the negative consequences of failing grades are quite serious and far outweigh any benefits. Even in upper grades, the fear of failure is a questionable motivation device. Better and more effective alternatives to failing grades need to be found, especially in a standards-based system. The use of I grades or incomplete grades present one meaningful alternative, especially if the necessary policies and resources are put in place to support those students who need additional assistance.

Policy 4: Using Zeros in Grading

Another related grading policy that hinders the implementation of standards-based reforms is the use of zeros. As part of their grading policies, many teachers assign zeros to students' work that is missed, neglected, or turned in late. However, the zero is seldom an accurate reflection of what a student has learned or is able to do (Raebeck 1993). Instead, zeros are typically assigned to punish students for not displaying appropriate effort or demonstrating adequate responsibility (Canady and Hotchkiss 1989; Stiggins and Duke 1991). If the grade is to represent how well students have learned or mastered established learning standards, then the practice of assigning zeros clearly misses the mark.

The effect of assigning zeros is intensified if combined with the practice of averaging to attain a student's overall course grade. Students readily see that receiving a single zero leaves them little chance for success because such an extreme score drastically skews the average. That is why, for example, in scoring Olympic events such as gymnastics, diving, or ice-skating; the highest and lowest scores are always eliminated. If they were not, one judge could control the entire competition simply by giving extreme scores.

Some teachers defend the practice of assigning zeros by arguing that they can not give students credit for work that is incomplete or not turned in—and that is certainly true. But there are far better ways to motivate and encourage students to complete assignments in a timely manner than through the use of zeros, especially considering the overwhelmingly negative effects.

Remedy

Students certainly should learn to accept responsibility for their actions and should be held accountable for their work. Nevertheless, no evidence demonstrates that assigning zeros helps teach students these lessons. Unless educators are willing to admit that grades are used to show evidence of students' lack of effort and responsibility, then alternatives to the practice of assigning zeros must be found.

One alternative approach is to assign an I (or Incomplete) grade with explicit requirements for completing the work, as addressed in the preceding discussion. For example, students whose work is incomplete or not turned in on time might be required to attend after-school study sessions or special Saturday classes until their work is completed to a satisfactory level. In other words, they are not "let off the hook" with a zero. Instead, students learn that they have certain responsibilities in school and that their actions have specific consequences. Not completing assigned work on time means that students must attend special after-school sessions to complete the work. Implementing such a policy may require additional funding and support; still, the payoffs are likely to be great. Not only is this approach more beneficial to students than simply assigning a zero, it is also a lot more fair. In addition, it helps make the grade a more accurate reflection of what students have learned.

Summary

To successfully implement standards-based reforms, educational leaders must take a broader and more systemic view of their efforts. Instead of focusing narrowly on curriculum and assessment issues, they must expand their perspective to consider organizational policies that can hinder success, especially in the area of grading and reporting student learning. Although grading will always be a process of professional judgment, making those judgments requires careful thought and continual reflection on the purpose of the activity. If grades are to represent information about the adequacy of students' achievement and performance with respect to clear learning standards, then the evidence used to determine grades must denote what students have learned and are able to do. To allow other factors to influence grades or marks misrepresents students' learning attainment.

Grading requires careful planning, thoughtful judgment, a clear focus on purpose, excellent communication skills, and an overriding concern for students. Such qualities are necessary to ensure grading policies and practices that provide high-quality information on student learning in any standards-based learning environment.

References

- + Adelman, C. 1999. Answers in the tool box: Academic intensity, attendance patterns, and Bachelor's degree attainment. Washington, D.C.: Office of Educational Research and Improvement, U.S. Department of Education.
 - + Bernetich, E. 1998, February 6. Personal communication to Thomas R. Guskey from Edward Bernetich, Principal of Beachwood Middle School, Beachwood, Ohio.
 - + Bloom, B. S. 1976. Human characteristics and school learning. New York: McGraw-Hill.
 - + Bloom, B. S., G. F. Madaus, and J. T. Hastings. 1981. Evaluation to improve learning. New York: McGraw-Hill.
 - + Bracey, G. W. 1994. Grade inflation? Phi Delta Kappan 76 (4): 328–29.
 - + Bracey, G. W. 1999. Getting that sheepskin. Phi Delta Kappan 81 (2): 169–70.
 - + Cameron, J., and W. D. Pierce. 1994. Reinforcement, reward, and intrinsic motivation: A meta-analysis. Review of Educational Research 64 (3): 363–423.
 - + ———. 1996. The debate about rewards and intrinsic motivation: Protests and accusations do not alter the results. Review of Educational Research 66 (1): 39–51.
 - + Canady, R. L., and P. R. Hotchkiss. 1989. It's a good score! Just a bad grade. Phi Delta Kappan 71 (1): 68–71.
 - + Chastain, K. 1990. Characteristics of graded and ungraded compositions. Modern Language Journal 74 (1): 10–14.
 - + Ebel, R. L. 1979. Essentials of educational measurement. 3d ed. Englewood Cliffs, N.J.: Prentice Hall.
 - + Engel, P. 1991. Tracking progress toward the school readiness goal. Educational Leadership 48 (5): 39–42.
 - + Feldmesser, R. A. 1971. The positive functions of grades. Paper presented at the annual meeting of the American Educational Research Association, New York.
 - + Gray, K. 1993. Why we will lose: Taylorism in America's high schools. Phi Delta Kappan 74 (5): 370–74.
 - + Haladyna, T. M. 1999. A complete guide to student grading. Boston: Allyn and Bacon.
 - + Johnson, D. W., and R. T. Johnson. 1989. Cooperation and competition: Theory and research. Endina, Minn.: Interaction.
 - + Johnson, D. W., L. Skon, and R. T. Johnson. 1980. Effects of cooperative, competitive, and individualistic conditions on children's problem-solving performance. American Educational Research Journal 17 (1): 83–93.
 - + Johnson, R. T., D. W. Johnson, and M. Tauer. 1979. The effects of cooperative, competitive, and individualistic goal structures on students' attitudes and achievement. Journal of Psychology 102: 191–98.
 - + Krumboltz, J. D., and C. J. Yeh. 1996. Competitive grading sabotages good teaching. Phi Delta Kappan 78 (4): 324–26.
 - + Kuehner, J. C. 1998. Bad news for bad grades: Beachwood school puts teeth in D's and F's by requiring extra classes or summer school. The Plain Dealer, 31 January, 1-A, 7-A.
 - + Lieberman, A., ed. 1995. The work of restructuring schools: Building from the ground up. New York: Teachers College Press.
 - + Raebeck, B. 1993. Exploding myths, exploring truths: Humane, productive grading and grouping in the quality middle school. Paper presented at the annual conference and exhibit of the National Middle School Association, Portland, Ore.
 - + Roderick, M., and E. Camburn. 1999. Risk and recovery from course failure in the early years of high school. American Educational Research Journal 36 (2): 303–43.
 - + Selby, D., and S. Murphy. 1992. Graded or degraded: Perceptions of letter-grading for mainstreamed learning-disabled students. British Columbia Journal of Special Education 16 (1): 92–104.
 - + Smith, D. 1999, July 2. Personal communication to Thomas R. Guskey from David Smith, Principal of West Springfield High School, Springfield, Va.
 - + Stiggins, R. J., and D. L. Duke. 1991. District grading policies and their potential impact on at-risk students. Paper presented at the annual meeting of the American Educational Research Association, 3-7 April, Chicago, Ill.
 - + Wood, L. A. 1994. An unintended impact of one grading practice. Urban Education 29 (2): 188–201.
- Note. This article is based on material included in Developing Grading and Reporting Systems for Student Learning, by T. R. Guskey and J. M. Bailey, to be published in 2001 by Corwin Press.*