

## The Dangerous Myth of Grade Inflation

ALFIE KOHN

Grade inflation got started . . . in the late '60s and early '70s. The grades that faculty members now give . . . deserve to be a scandal.

—*Professor Harvey Mansfield, Harvard University, 2001*

Grades A and B are sometimes given too readily—Grade A for work of no very high merit, and Grade B for work not far above mediocrity. . . . One of the chief obstacles to raising the standards of the degree is the readiness with which insincere students gain passable grades by sham work.

—*Report of the Committee on Raising the Standard,  
Harvard University, 1894*

Complaints about grade inflation have been around for a very long time. Every so often a fresh flurry of publicity pushes the issue to the foreground again, the latest example being a series of articles in the *Boston Globe* in 2001 that disclosed—in a tone normally reserved for the discovery of entrenched corruption in state government—that a lot of students at Harvard were receiving As and being graduated with honors.

The fact that people were offering the same complaints more than a century ago puts the latest bout of harrumphing in perspective, not unlike those quotations about the disgraceful values of the younger generation that turn out to be hundreds of years old. The long history of indignation also pretty well derails any attempts to place the blame for higher grades on a residue of bleeding-heart liberal professors hired in the '60s. (Unless, of course, there was a similar countercultural phenomenon in the 1860s.)

Yet on campuses across America today, academe's usual requirements for supporting data and reasoned analysis have been suspended for some reason where this issue is concerned. It is largely accepted on faith that grade inflation—an upward shift in students' grade point averages without a similar rise in achievement—exists, and that it is a bad thing. Meanwhile, the truly substantive issues surrounding grades and motivation have been obscured or ignored.

The fact is that it is hard to substantiate even the simple claim that grades have been rising. Depending on the time period we are talking about, that claim may well be false. In their book *When Hope and Fear Collide*, Arthur Levine and Jeanette Cureton tell us that more undergraduates in 1993 reported receiving As (and fewer reported receiving grades of C or below) compared to their counterparts in 1969 and 1976 surveys.<sup>1</sup> Unfortunately, self-reports are notoriously unreliable, and the numbers become even more dubious when only a self-selected, and possibly an unrepresentative, segment bothers to return the questionnaires. (One out of three failed to do so in 1993; no information is offered about the return rates in the earlier surveys.)

To get a more accurate picture of whether grades have changed over the years, one needs to look at official student transcripts. Clifford Adelman, a senior research analyst with the U.S. Department of Education, did just that, reviewing transcripts from more than 3,000 institutions and reporting his results in 1995. His finding: "Contrary to the widespread lamentations, grades actually declined slightly in the last two decades." Moreover, a report released in 2002 by the National Center for Education Statistics revealed that fully 33.5 percent of American undergraduates had a grade point average of C or below in 1999–2000, a number that ought to quiet "all the furor over grade inflation," according to a spokesperson for the Association of American Colleges and Universities. (A review of other research suggests a comparable lack of support for claims of grade inflation at the high school level.)<sup>2</sup>

However, even where grades *are* higher now compared to then, that does not constitute proof that they are inflated. The burden rests with critics to demonstrate that those higher grades are undeserved, and one can cite any number of alternative explanations. Maybe students are turning in better assignments. Maybe instructors used to be too stingy with their marks and have become more reasonable. Maybe the concept of assessment itself has evolved, so that today it is more a means for allowing students to demonstrate what they know rather

than for sorting them or “catching them out.” (The real question, then, is why we spent so many years trying to make good students look bad.) Maybe students are not forced to take as many courses outside their primary areas of interest in which they did not fare as well. Maybe struggling students are now able to withdraw from a course before a poor grade appears on their transcripts. (Say what you will about that practice, it challenges the hypothesis that the grades students receive in the courses they complete are inflated.)

The bottom line: No one has ever demonstrated that students today get As for the same work that used to receive Bs or Cs. We simply do not have the data to support such a claim. Consider the most recent, determined effort by a serious source to prove that grades are inflated: “Evaluation and the Academy: Are We Doing the Right Thing?” a report released in 2002 by the American Academy of Arts and Sciences.<sup>3</sup> Its senior author is Henry Rosovsky, formerly Harvard’s dean of the faculty. The first argument offered in support of the proposition that students could not possibly deserve higher grades is that SAT (Scholastic Aptitude Test) scores have dropped during the same period that grades are supposed to have risen. But this is a patently inapt comparison, if only because the SAT is deeply flawed. It has never been much good even at predicting grades during the freshman year in college, to say nothing of more important academic outcomes. A four-year analysis of almost 78,000 University of California (UC) students, published in 2001 by the UC president’s office, found that the test predicted only 13.3 percent of variation in freshman grades, a figure roughly consistent with hundreds of previous studies.<sup>4</sup>

Even if one believes that the SAT is a valid and valuable exam, however, the claim that scores are dropping is a poor basis for the assertion that grades are too high. First, it is difficult to argue that a standardized test taken in high school and grades for college course work are measuring the same thing.

Second, changes in aggregate SAT scores mostly reflect the proportion of the eligible population that has chosen to take the test. The American Academy’s report states that average SAT scores dropped slightly from 1969 to 1993. But over that period, the pool of test takers grew from about one-third to more than two-fifths of high school graduates—an addition of more than 200,000 students.

Third, a decline in overall SAT scores is hardly the right benchmark against which to measure the grades earned at Harvard or other elite

institutions. Every bit of evidence I could find—including a review of the SAT scores of entering students at Harvard over the past two decades, at the nation’s most selective colleges over three and even four decades, and at all private colleges since 1985—uniformly confirms a virtually linear rise in both verbal and math scores, even after correcting for the renorming of the test in the mid-1990s. To cite just one example, the latest edition of “Trends in College Admissions” reports that the average verbal-SAT score of students enrolled in all private colleges rose from 543 in 1985 to 558 in 1999. Thus those who regard SAT results as a basis for comparison should *expect* to see higher grades now rather than assume that they are inflated.

The other two arguments made by the authors of the American Academy’s report rely on a similar sleight of hand. They note that more college students are now forced to take remedial courses but offer no reason to think that this is especially true of the relevant student population—namely, those at the most selective colleges who are now receiving As instead of Bs.<sup>5</sup>

Finally, they report that more states are adding high school graduation tests and even standardized exams for admission to public universities. Yet that trend can be explained by political factors and offers no evidence of an objective decline in students’ proficiency. For instance, scores on the National Assessment of Educational Progress, known as “the nation’s report card” on elementary and secondary schooling, have shown very little change over the past couple of decades, and most of the change that has occurred has been for the better. As David Berliner and Bruce Biddle put it in their tellingly titled book *The Manufactured Crisis*,<sup>6</sup> the data demonstrate that “today’s students are at least as well informed as students in previous generations.” The latest round of public school bashing—and concomitant reliance on high-stakes testing—began with the Reagan administration’s “Nation at Risk” report, featuring claims now widely viewed by researchers as exaggerated and misleading.

Beyond the absence of good evidence, the debate over grade inflation brings up knotty epistemological problems. To say that grades are not merely rising but inflated—and that they are consequently “less accurate” now, as the American Academy’s report puts it — is to postulate the existence of an objectively correct evaluation of what a student (or an essay) deserves, the true grade that ought to be uncovered and

honestly reported. It would be an understatement to say that this reflects a simplistic and an outdated view of knowledge and learning.

In fact, what is most remarkable is how rarely learning even figures into the discussion. The dominant disciplinary sensibility in commentaries on this topic is not that of education—an exploration of pedagogy or assessment—but rather of economics. That is clear from the very term “grade inflation,” which is, of course, just a metaphor. Our understanding is necessarily limited if we confine ourselves to the vocabulary of inputs and outputs, incentives, resource distribution, and compensation.

Suppose, for the sake of the argument, we assumed the very worst—not only that students are getting better grades than did their counterparts of an earlier generation, but that the grades are too high. What does that mean, and why does it upset some people so?

To understand grade inflation in its proper context, we must acknowledge a truth that is rarely named: The crusade against it is led by conservative individuals and organizations who regard it as analogous—or even related—to such favorite whipping boys as multicultural education, the alleged radicalism of academe, “political correctness” (a label that permits the denigration of anything one does not like without having to offer a reasoned objection), and too much concern about students’ self-esteem. Mainstream media outlets and college administrators have allowed themselves to be put on the defensive by accusations about grade inflation, as can be witnessed when deans at Harvard plead *nolo contendere* and dutifully tighten their grading policies.

What are the critics assuming about the nature of students’ motivation to learn, about the purpose of evaluation and of education itself? (It is surely revealing when someone reserves time and energy to complain bitterly about how many students are getting As—as opposed to expressing concern about, say, how many students have been trained to think that the point of going to school is to get As.)

“In a healthy university, it would not be necessary to say what is wrong with grade inflation,” Harvey Mansfield asserted in an opinion article in 2001 in the *Chronicle of Higher Education*.<sup>7</sup> That, to put it gently, is a novel view of health. It seems reasonable to expect those making an argument to be prepared to defend it, and also valuable to bring their hidden premises to light. The assumptions that follow seem to underlie the grave warnings about grade inflation:

*The professor's job is to sort students for employers or graduate schools.* Some are disturbed by grade inflation—or, more accurately, grade compression—because it then becomes harder to spread out students on a continuum, ranking them against one another for the benefit of postcollege constituencies. One professor asks, by way of analogy, “Why would anyone subscribe to *Consumers Digest* if every blender were rated a ‘best buy?’”

But how appropriate is such a marketplace analogy? Is the professor's job to rate students like blenders for the convenience of corporations, or to offer feedback that will help students learn more skillfully and enthusiastically? (Notice, moreover, that even consumer magazines do not grade on a curve. They report the happy news if it turns out that every blender meets a reasonable set of performance criteria.)

Furthermore, the student-as-appliance approach assumes that grades provide useful information to those postcollege constituencies. Yet growing evidence—most recently in the fields of medicine and law, as cited in publications such as the *Journal of the American Medical Association* and the *American Educational Research Journal*—suggests that grades and test scores do not in fact predict career success, or much of anything beyond subsequent grades and test scores.

*Students should be set against one another in a race for artificially scarce rewards.* “The essence of grading is exclusiveness,” Mansfield said in one interview. Students “should have to compete with each other,” he said in another. A chemistry professor at the University of Wisconsin at La Crosse said, “We cannot have half our students at the head of the class.” In other words, even when no graduate school admissions committee pushes for students to be sorted, they ought to be sorted anyway, with grades reflecting relative standing rather than absolute accomplishment. In effect, this means that the game should be rigged so that no matter how well students do, only a few can get As. The question guiding evaluation in such a classroom is not “How well are they learning?” but “Who is beating whom?” The ultimate purpose of good colleges, this view holds, is not to maximize success but to ensure that there will always be losers.

A bell curve may sometimes—but only sometimes—describe the range of knowledge in a roomful of students at the beginning of a course. When it is over, though, any responsible educator hopes that the results would skew drastically to the right, meaning that most students learned what they had not known before. Thus in their important

study *Making Sense of College Grades*, Ohmer Milton, Howard Pollio, and James Eison write, “It is not a symbol of rigor to have grades fall into a ‘normal’ distribution; rather, it is a symbol of failure—failure to teach well, failure to test well, and failure to have any influence at all on the intellectual lives of students.”<sup>8</sup> Making sure that students are continually re-sorted, with excellence turned into an artificially scarce commodity, is almost perverse. Excellence is not the same thing as victory.

What does relative success signal about student performance in any case? The number of peers that a student has bested tells us little about how much she knows and is able to do. Moreover, such grading policies may create a competitive climate that is counterproductive for winners and losers alike, to the extent that it discourages a free exchange of ideas and a sense of community that is conducive to exploration.

*Harder is better (or higher grades mean lower standards)*. Compounding the tendency to confuse excellence with victory is a tendency to confuse quality with difficulty—as evidenced in the accountability fad that has elementary and secondary education in its grip just now, with relentless talk of “rigor” and “raising the bar.” The same confusion shows up in higher education when professors pride themselves not on the intellectual depth and value of their classes but merely on how much reading they assign, how hard their tests are, how rarely they award good grades, and so on. “You’re going to have to *work* in here!” they announce, with more than a hint of machismo and self-congratulation.

Some people might defend that posture on the grounds that students will perform better if As are harder to come by. In fact, the evidence on this question is decidedly mixed. Stringent grading sometimes has been shown to boost short-term retention as measured by multiple-choice exams—never to improve understanding or promote interest in learning. An analysis, released in 2000 by Julian R. Betts and Jeff Grogger, professors of economics at the University of California at San Diego and at Los Angeles, respectively, found that tougher grading was initially correlated with higher test scores. But the long-term effects were negligible—with the exception of minority students, for whom the effects were negative.

It appears that something more than an empirical hypothesis is behind the “harder is better” credo, particularly when it is set up as a painfully false dichotomy: Those easy-grading professors are too lazy to care, or too worried about how students will evaluate them, or overly concerned about their students’ self-esteem, whereas *we* are the last

defenders of what used to matter in the good old days. High standards! Intellectual honesty! No free lunch!

The American Academy's report laments an absence of "candor" about this issue. Let us be candid then. Those who grumble about undeserved grades sometimes exude a cranky impatience with—or even contempt for—the late adolescents and young adults who sit in their classrooms. Many people teaching in higher education, after all, see themselves primarily as researchers and regard teaching as an occupational hazard, something they are not very good at, were never trained for, and would rather avoid. It would be interesting to examine the correlation between one's view of teaching (or of students) and the intensity of one's feelings about grade inflation. Someone also might want to examine the personality profiles of those who become infuriated over the possibility that someone, somewhere, got an A without having earned it.

*Grades motivate.* With the exception of orthodox behaviorists, psychologists have come to realize that people can exhibit qualitatively different kinds of motivation: intrinsic, in which the task itself is seen as valuable, and extrinsic, in which the task is just a means to the end of gaining a reward or escaping a punishment. The two are not only distinct but often inversely related. Scores of studies have demonstrated, for example, that the more people are rewarded, the more they come to lose interest in whatever had to be done in order to get the reward. (That conclusion is essentially reaffirmed by the latest major meta-analysis on the topic: a review of 128 studies, published in 1999 by Edward L. Deci, Richard Koestner, and Richard M. Ryan.)<sup>9</sup>

Those unfamiliar with that basic distinction, let alone the supporting research, may be forgiven for pondering how to "motivate" students, then concluding that grades are often a good way of doing so, and consequently worrying about the impact of inflated grades. But the reality is that it does not matter how motivated students are; what matters is *how* students are motivated. A focus on grades creates, or at least perpetuates, an extrinsic orientation that is likely to undermine the love of learning we are presumably seeking to promote.

Three robust findings emerge from the empirical literature on the subject: Students who are given grades, or for whom grades are made particularly salient, tend to display less interest in what they are doing, fare worse on meaningful measures of learning, and avoid more challenging tasks when given the opportunity—as compared to those in a

nongraded comparison group. College instructors cannot help noticing, and presumably being disturbed by, such consequences, but they may lapse into blaming students (“grade grubbers”) rather than understanding the systemic sources of the problem. A focus on whether too many students are getting As suggests a tacit endorsement of grades that predictably produces just such a mind-set in students.

These fundamental questions are almost completely absent from discussions of grade inflation. The American Academy’s report takes exactly one sentence—with no citations—to dismiss the argument that “lowering the anxiety over grades leads to better learning,” ignoring the fact that much more is involved than anxiety. It is a matter of why a student learns, not only how much stress he feels. Nor is the point just that low grades hurt some students’ feelings, but that grades, per se, hurt all students’ engagement with learning. The meaningful contrast is not between an A and a B or C but between an extrinsic and an intrinsic focus.

Precisely because that is true, a reconsideration of grade inflation leads us to explore alternatives to our (often unreflective) use of grades. Narrative comments and other ways by which faculty members can communicate their evaluations can be far more informative than letter or number grades, and much less destructive. Indeed, some colleges—for example, Hampshire, Evergreen State, Alverno, and New College of Florida—have eliminated grades entirely, as a critical step toward *raising* intellectual standards. Even the American Academy’s report acknowledges that “relatively undifferentiated course grading has been a traditional practice in many graduate schools for a very long time.” Has that policy produced lower-quality teaching and learning? Quite the contrary: Many people say they did not begin to explore ideas deeply and passionately until graduate school began and the importance of grades diminished significantly.

If the continued use of grades rests on nothing more than tradition (“We’ve always done it that way”), a faulty understanding of motivation, or excessive deference to graduate school admissions committees, then it may be time to balance those factors against the demonstrated harms of getting students to chase As. Milton and his colleagues discovered—and others have confirmed—that a “grade orientation” and a “learning orientation” on the part of students tend to be inversely related. That raises the disturbing possibility that some colleges are institutions of higher learning in name only, because the paramount

question for students is not “What does this mean?” but “Do we have to know this?”

A grade-oriented student body is an invitation for the administration and faculty to ask hard questions: What unexamined assumptions keep traditional grading in place? What forms of assessment might be less destructive? How can professors minimize the salience of grades in their classrooms so long as grades must still be given? And, if the artificial inducement of grades disappeared, then what sort of teaching strategies might elicit authentic interest in a course?

To engage in this sort of inquiry, to observe real classrooms, and to review the relevant research is to arrive at one overriding conclusion: The real threat to excellence is not grade inflation at all; it is grades.<sup>10</sup>

## NOTES

1. Arthur Levine and Jeanette Cureton, *When Hope and Fear Collide* (San Francisco: Jossey-Bass, 1998).

2. A subsequent analysis by Adelman, which reviewed college transcripts from students who were graduated from high school in 1972, 1982, and 1992, confirmed that there was no significant or linear increase in average grades over that period. The average GPA for those three cohorts was 2.70, 2.66, and 2.74, respectively. The proportion of As and Bs received by students was 58.5 percent in the '70s, 58.9 percent in the '80s, and 58.0 percent in the '90s. Even when Adelman looked at “highly selective” institutions, he again found very little change in average GPA over the decades.

3. Henry Rosovsky and Matthew Hartley, “Evaluation and the Academy: Are We Doing the Right Thing? Grade Inflation and Letters of Recommendation,” Occasional Paper (Cambridge, MA: American Academy of Arts and Sciences, 2002).

4. I outlined numerous other problems with the test in “Two Cheers for an End to the SAT,” *The Chronicle*, March 9, 2001, available at [www.alfiekohn.org](http://www.alfiekohn.org).

5. Adelman’s newer data challenge the premise that there has been *any* increase. In fact, “the proportion of all students who took at least one remedial course (in college) dropped from 51 percent in the [high school] class of 1982 to 42 percent in the class of 1992.”

6. David Berliner and Bruce Biddle, *The Manufactured Crisis* (Reading, MA.: Addison-Wesley, 1995).

7. Harvey Mansfield, “Grade Inflation: It’s Time to Face the Facts,” *Chronicle of Higher Education*, April 6, 2001.

8. Ohmer Milton, Howard Pollio, and James Eison, *Making Sense of College Grades* (San Francisco: Jossey-Bass, 1986).

9. Edward L. Deci, Richard Koestner and Richard M. Ryan, "A Meta-analytic Review of Experiments Examining the Effects of Extrinsic Rewards on Intrinsic Motivation," *Psychological Bulletin* 125, (1999): 627–68.

10. This article is based on one that first appeared in *Chronicle of Higher Education* 49:11 (November 8, 2002) and Mr. Kohn's talk at the "Academic Standards and Grade Inflation Conference," October 11, 2003.