

## **“Common Sense” Grading Policy Recommendations? A Rebuttal of Bulk and Monte**

L. L. Nelson

### **Introduction: Statement of the Controversial Problem**

Bulk and Monte begin by reporting on the current grade inflation situation at Harvard and assert that “graduating with honors is no longer much of a distinction at Harvard or any other grade inflated college” (1:1). Yes, Harvard’s lack of distinction is what justifies its low tuition and causes its student and faculty recruitment problem, as well as its shabby endowments. They report that Harvard’s solution to the problem has been to limit honors to 55% of its graduates and to adopt a conventional “A through F” grading scale. I am willing to bet a great deal of money that the average grade at Harvard will be higher than a “gentleman’s C” after these steps. Bilby notes in his position paper that the grade inflation problem is most pronounced at elite institutions such as Harvard, Yale, and Duke. This is a false analogy to the situation at UW-L.

By the way, a "gentleman's C" is a phrase that originated at Oxford and Cambridge and meant a "C" grade given to a gentleman for doing very little work in a course. One could not, of course, embarrass a member of the nobility or gentry with a "D" or an "F." A parallel phrase, "gentleman commoner" was used to denote a commoner enrolled at university receiving special privileges of some kind. Most students at Oxford and Cambridge were members of the nobility or landed gentry when these phrases originated long before the 20th century. Grade inflation is not a new problem.

Bulk and Monte then assert that we have a grade inflation problem at UW-L and support their assertion with a blatant slippery slope fallacy. They report that between 1976 and 1988 the average GPA at UW-L held steady at 2.84. Between 1989 and 2002, the average GPA increased to 3.20. They extrapolate that if the trend continues, by 2025 the average GPA will be 3.70. Their assumption that the trend will continue is alarmist. It is anyone’s guess what the trend will be over the next two or three decades. Over the last 27 years, there has been an increase of .36 or 9% of the finite 4.00 scale. Symbolically it represents a shift from slightly below a B average to slightly above a B average. On a 100 point scale, it is the equivalent of an increase from a 71 to an 80. In a 15 credit load of 3 credit courses, it is a change from four B’s and a C, slightly below a B average, to four B’s and an A, slightly above a B average. Feeling a little less panicked?

**I agree with Bulk and Monte that a moderate inflationary trend exists at UW-L and is a cause for some concern. I do not believe we are in a grade inflation crisis. However, grade inflation is an issue we need to address, as do all institutions of higher education. Bulke and Monte assert that 1) grades at UW-L “are becoming so highly inflated that they are beginning to lose all value and meaning” and 2) “extreme lack of consistency in grading practices, both across and within, departments creates a situation that is unfair to our students” (1:3). If this were the case, we would be in sad straits indeed. Do they make their case?**

## **A Brief and Highly Anecdotal History**

They assert that during the 1960's "it was normative . . . (at least among the more prestigious colleges of the nation) to grade on a bell shaped curve." They further assert that this prevented problems of grade inflation because it assured that the average GPA would "hover around 2.00, the so-called 'gentleman's C'" (1:4). On the contrary, there was a grade inflation problem in the 1960's and 1970's driven by GPA requirements of student deferments during the Viet Nam War. According to Bulk and Monte, college during the 1960's was "more of an elitist institution with a relatively wealthier group of Americans attending" (2:1). It is certainly true that the numbers of Americans attending college increased tremendously over the course of the 20<sup>th</sup> century, but the shift away from elitist institutions attended only by the wealthy was underway well before the 1960's: G.I. benefits were used by many returning servicemen after WWII. Bulk and Monte then tell us, "Today, a much broader class spectrum attends college, a fact that has resulted in a drawing down of the national mean SAT scores." They appear to claim that it is this shift in the economic demographics of college students that has caused SAT scores to fall and be re-normed. If so, Bulk and Monte have confused cause and effect. The drop in SAT scores was due to multiple plausible significant causes, only one of which was that more people from a broader range of economic backgrounds were taking the exam. They have also compressed more than half a century of higher education history.

Bulk and Monte state "it is a truism" that the distribution of grades "is a fundamentally arbitrary matter to be determined entirely by the discretion of the course instructor" (2:2). According to Bulk and Monte, the grade distribution bears no necessary or fixed relationship to "how intelligent the students being graded are, how frequently they attend class, how well they perform on exams, or anything else" (2:2). While this may be their attempt to acknowledge the autonomous authority of instructors to assign grades, most conscientious instructors would strongly disagree. In fact, many would argue that good teaching makes grade distributions far from arbitrary. There is nothing arbitrary about identifying the goals and objectives of a course, designing assignments and exams to achieve those objectives with students, setting up and communicating clearly defined expectations and standards for the students, and then evaluating their performance in achieving the goals and objectives relative to these established expectations and standards.

**In the same paragraph, Bulk and Monte would have us believe that "until recent times, few college instructors ever entertained the realistic possibility that normative college grading distributions would contract so much that the C and lower grades would practically disappear from the grading radar screen" (2:2). On the contrary, for as long as I have worked at UW-L, grading policies have insured that this will be the case.**

First, since I began teaching here in 1983, students have been able to erase D and F grades by retaking the courses in which they performed poorly. For continuing UW-L students, a D or F is only a temporary problem that will be calculated out of their GPAs when they take the course again and a second, higher grade is substituted for the D or F they originally earned. This is fundamentally inflationary. Secondly, students are able to drop courses until a week after midterm. Some students will drop a course in which they are receiving a C because the C is the lowest grade that will stubbornly remain in the calculation of their GPAs. Finally, to remain in good academic standing and continue to attend UW-L, students must maintain a 2.00 or C average, not a D average. According to Bulk and Monte, a C is defined as “satisfactory or average work.” In essence, the 2.00 GPA benchmark for good academic standing means that we demand a minimum of “satisfactory or average work” from our students. Students who consistently perform at a “barely satisfactory or poor” level do not remain in good standing at UW-L. Their GPAs disappear from the calculation of the average GPA of continuing students. Since we demand the minimum of satisfactory or average performance to continue at UW-L, all UW-L students do have to be “above average” by definition.

I do not doubt that “grade rage” is a more frequent phenomenon these days. Bulk and Monte describe students who attack instructors when they receive a B or less, ostensibly because students are used to receiving and/or believe they are entitled to higher grades. In my own experience, this is more likely to occur with freshmen and transfer students than it is with continuing UW-L students. Another cause of “grade rage” is the use of highly selective admission requirements for some majors with extremely limited seats into which many students want to be accepted. For these students, any deviation from a 4.00 may seriously disadvantage them in future admission to the desired program. Although their behavior is not acceptable, their distress is understandable.

### **Does Grade Inflation Matter?**

In this section of their position paper, the authors explain why grade inflation is a problem. It is a problem because giving every student an A “would not fairly represent the actual student performances exhibited” (3:3). They argue that since all students are different, they must all earn different grades. They enter our classrooms with different backgrounds and skill levels; therefore, they must leave the same way. When all students receive A’s, we lie and bankrupt our professional credibility.

I agree that unfair grading damages our professional credibility. The people with whom it damages us first and most importantly are students. But it is very important to remember that different disciplines grade fairly in different ways. What are fair methods of grading in some courses and disciplines may not be fair, or even applicable, in others.

Here the mean-spirited nature of this position paper most clearly reveals itself. Bulk and Monte state, "It is the faculty member's foremost responsibility to award course grades that reflect the relative degree of student mastery and achievement of course objectives." Few if any faculty would disagree. However, they go on to assert "if this cannot be done then there must be something wrong with the course (i.e. it is not sufficiently challenging) or with the instructor's competence as it relates to grading ability" (4:2). In other words, they place the blame for the asserted problem of grade inflation squarely on the backs of fellow faculty who, if they are engaging in "high end" grading, must necessarily be incompetent in either course design and/or ability to discriminate among student performances. They assert that it is highly unlikely that all the students in a given class can achieve "top end" course mastery. They reassert the argument that since all students are different, there must be differences in the grades they earn. However, at this point they also shift ground from the individual instructors to the institution as a whole and point out that grade inflation is in the aggregate of grades given rather than being a function of "a few isolated courses or instructors." The tar is being spread rather thickly and widely here.

Grading practices and grade inflation are affected by interrelationships among instructors and courses and the aggregated number itself and the grading system and the students and the institutional culture as well as the societal context in which that institution exists. Grading is undeniably a complex symbolic action; consequently, there are many ways in which the assertions made above can be refuted.

First it is most important to reply that these are NOT the only two ways in which all students in a class might receive "high end" grades. This is a classic example of the fallacy of the excluded middle. This either/or thinking excludes a myriad of possibilities, not least of which is that the educational experience worked well and the students learned what they needed to learn. It would be gratifying to think that some small degree of grade inflation might be due to improvement in instruction based on assessment efforts.

It is also important to stress that students in our classes are not random samples representative of all possible students but are instead comprised of groups structured by academic programs and self-selected to varying degrees. We begin "de-randomizing" students in the admissions process; not even the incoming freshman classes are random samples. Bilby discusses the positive change in the qualifications of UW-L students as admission standards have become more rigorous and relates these changes to average GPAs. Even in the General Education program, students self-select into different courses for many different reasons. And of course, few students select majors and minors in which they will perform poorly, have no interest and/or no motivation to succeed.

Finally, differences among students entering our classes does lead to different performances in our courses, but many of us have also worked harder in recent years to more fairly accommodate many of these differences—from learning styles to sex and gender to ethnic and racial heritage to disabilities—and sometimes even to celebrate

them. It would be gratifying to think that some slight degree of grade inflation might be due to our doing a better job at providing courses in which both the content and pedagogy are appropriately adapted to students, not just in the featureless aggregate, but also in their diversity.

Bulk and Monte describe a worst case scenario, in which "not just a few isolated . . . instructors" but a significant number of instructors are giving students high end grades without bothering to discriminate among levels of student performance, abdicating their professional responsibility to society at large and seriously damaging our professional credibility. I do not deny that there are instructors who engage in the indiscriminate awarding of high end grades or that there are courses in which the performance bar is set at a level where most of our students can meet or exceed the requirements for a B grade. There are also instructors who engage in the indiscriminate awarding of low end grades and there are courses in which many of our students have difficulty earning a B. However, I believe that most of my colleagues are conscientious about the teaching/learning process, and the level at which the performance bar is set must be related to the content, purposes, and objectives of specific courses.

The last significant issue raised with regard to the fairness issue is that, according to Bulk and Monte, students in courses and programs which are less grade inflated are disadvantaged in relation to other students in competition for scholarships and honors. In order to accept this assertion, one has to accept the premise that the majority of courses and programs are indeed grade inflated to a significant degree and that a minority of others are not.

An alternative explanation also exists: a minority of courses and programs are poorly taught, badly adapted to their student audiences, and more focused on producing classic bell curve results in grading than in fostering student learning. Students who are take these courses are definitely disadvantaged in the competition for scholarships and honors.

Certain programs on this campus have complained for years that our students are poorly prepared and have justified "low end" grading on this basis. However, on standardized tests, Wisconsin high school students perform among the best in the country, and we are enrolling the top 25% of those students by class rank. Since the quality of the student body has been significantly improved as entrance requirements have become more selective, it may be time for programs to reconsider course design, curriculum design, pedagogical methods, and grading premises. As entrance requirements have become more selective, justification for "low end" grading has been proportionately weakened. For example, it is highly inappropriate when a General Education course has a D-F-drop rate of 30-40% and when instructors lose large numbers of students between the 10<sup>th</sup> day enrollment figures and the week after midterm.

## **A Critique of Proposed Solutions and an Alternative Proposal for Addressing Grade Inflation at UW-L**

Bulk and Monte set up the "straw man" of eliminating grades entirely and reject this proposal as impractical, which it certainly is (7:6-8:1). They then propose that the mean grade earned/awarded in each class be provided on student transcripts (8:2-9:1). Finally, they propose that the mean grades earned/awarded in courses be used in evaluating faculty (9:2). They propose that there should be no "binding compulsion" for every class to conform to a B average, but that there should be a "forceful expectation" for faculty to generally adhere to a university-wide standard B average. If such a policy does not produce "voluntary compliance" then institution of "meaningful sanctions" could be necessary, since "it is an unfortunate fact of life that any policy, of any kind, that exists wholly without the possibility of meaningful sanctions is a policy likely to be ignored" (9:3).

Reporting the mean grade given in a course on student transcripts does not directly address the issue of grade inflation, though it may make it more noticeable to our publics. Certainly if students themselves want this information shared with prospective employers and graduate schools, it should be made available. If employers and graduate schools want this information, it should be made available. However, it should not be made available in order to address grade inflation, because it is not a solution to that problem. It is only a report of the problem.

Use of grading practices as a factor in evaluating faculty is standard, status quo practice, already being used to evaluate faculty in many departments. If the information were made more easily available it would be more used. Its relative importance compared to SEIs, student written comments, course design and management, peer observations, and other aspects of teaching evaluation is a matter for more specific debate among faculty in departments, colleges, and the Joint Promotion Committee.

Under no circumstances should we institute punitive measures specifically for failing to adhere to a certain average instructor GPA. This sort of precedent is a serious weakening of instructor prerogative to assign grades. This sort of policy does much more damage to our professional credibility than the current moderate inflationary trend at this institution. Such a policy says that we do not trust one another and the administration does not trust us to assign appropriate grades in our courses. In short, the harms of such a "solution" outweigh its benefits.

This "solution" is also redundant. Retention, merit, tenure, and promotion processes provide excellent vehicles for us as a faculty to use grading practices as part of peer evaluation and exert peer pressure with regard to inappropriate "high end" and "low end" grading. It only requires that we have information about grading practices readily available and that we weight this factor more significantly in our evaluation of instructional effectiveness. As these processes have important practical consequences for all of us, additional punitive consequences are redundant.

## Alternative Proposal for Addressing Grade Inflation at UW-L

I believe that a somewhat different approach should be used at UW-L to address the issue of grade inflation, in which we correct inflationary aspects of the grading *system* as well as change the university culture with respect to peer evaluation of faculty performance. Two steps should be taken.

1. CAPS should abandon the policy of erasing D and F grades earned in courses by students who then retake those courses for higher subsequent grades. D or F grades should be retained in the calculation of student GPAs permanently. Students should still be allowed to retake those courses for mastery of course material, but the initial grade earned should remain in the calculation of their GPAs and be balanced by the subsequent grade earned. Currently, a student who earns an F initially and subsequently earns an A has an A average. Under the proposed policy, the student would have a C average. This would necessarily result in a drop in the university's aggregate student GPA. This policy might be moderated by allowing freshmen and transfer students to "erase" D's and F's earned in their first or first two semesters enrolled at UW-L, just as we allow freshmen to drop courses without penalty with regard to the drop policy. This would allow new students to adjust to the standards of performance at UW-L before being held accountable.

The current policy wastes valuable seats in courses. The current policy encourages students performing at a C level in a course to give up and accept a D or F, knowing that the D or F will do less damage to their GPAs in the long run than the C they might have earned with sustained effort the first time around. The proposed policy would bring the C grade back "on the screen" in the range of grades earned by students. The C would become preferred to a D or F in terms of its impact on GPA for students. This also provides some motivation for students to invest more serious and sustained effort in courses the first time around.

2. All faculty and instructional academic staff, not just those eligible for promotion, should receive teaching assignment information forms every year and those forms should be used in departmental merit, retention, and tenure evaluations as well as promotion evaluations. This would enable more balanced and thorough peer evaluations by colleagues. Such a change would have a number of significant benefits. First, instructor GPAs would be established as a check and balance to instructor SEIs. This would help to discourage instructors from pandering to students by giving inflated grades in order to receive inflated SEI scores. It would allow us to provide more accurate and effective peer evaluation as well as better mentoring to faculty and instructional academic staff who need to adjust their performances. It would provide more consistency in the various levels of faculty evaluation. Most importantly, it would provide more detailed feedback to instructors about their grading practices, feedback that they do not now receive in detail unless they are eligible for promotion.

## Summary

Bulke and Monte's position paper is called "Common Sense Grading Policy Recommendations." An apology is due to poor Tom Paine. Bulke and Monte do not make their case that that 1) grades at UW-L "are becoming so highly inflated that they are beginning to lose all value and meaning" and 2) "extreme lack of consistency in grading practices, both across and within, departments creates a situation that is unfair to our students" (1:3). I agree that we have a moderate inflationary trend that we should address as a faculty. Solutions proposed by Bulk and Monte do not address the problem, are redundant, and/or have significant harms. An alternative and more measured approach would address inflationary policy in the grading system and put increased weight on grade distributions in the evaluations we already conduct. This would require that grade distribution information be made readily available to all faculty and instructional academic staff. Since we distribute this information to a limited group of faculty every year, it should pose no problem to make it available to all.

