

# Grading Policy Recommendations for UW-L

by

Jac D. Bulk (CLS) and Aaron Monte (CSAH)

## Introduction: Statement of the Problem

This position paper on undergraduate college grading policy is intended to stimulate critical thinking and discourse on a subject that is central to all of us as college educators. It is also intended to provoke some deep thought about what is *ultimately* fair and in our student's and university's best interest with respect to the grades we assign each semester. The stimulus for raising this issue at this particular time has been the unprecedented rise in grade inflation across the nation's colleges and the subsequent public alarm that has resulted from some recent revelations from academia. For example, *over 90 percent* of the Harvard student body graduated with honors during the 2001-2002 academic year. Clearly, not graduating with honors becomes a social stigma under these circumstances. Likewise, graduating with honors is no longer much of a distinction, at Harvard, or at any other grade-inflated institution. Not surprisingly, following the widespread criticism of this situation last spring, the Harvard faculty moved quickly to establish policies that now limit each of their honors categories to fixed percentages of students (5% *summa*, 20% *summa plus magna*, and 30% *cum laude*) and have done away with their idiosyncratic 15-point grading scale in favor of the more conventional "A through F" scale, where an **A** equals 4.00 points and an **F** 0.00 points in calculating student grade point averages (GPAs). These changes were supported by a unanimous voice vote of over 200 of Harvard's faculty and are now in effect for current freshmen. (*New York Times*, May 29, 2002)

With respect to grade inflation in general, the University of Wisconsin–La Crosse has followed the national trend. Institutional records at UW-La Crosse clearly show that the mean undergraduate GPA remained on a plateau of about 2.84 from the spring semester of 1976 through the spring semester of 1988. However, from 1989 to the present, the UW-L mean GPA has ratcheted steadily upward to its present level of 3.20 and continues to climb every year (see attached **figure**). In fact, extrapolating the grade distribution data since 1975 and using the trend of the past decade, we find that by the year 2025, the mean UW-L GPA would become a 3.7 (or an **AB**). It is important to keep in mind that these figures do include a handful of departments whose mean GPAs are well below the overall university mean GPA. If we were to remove these "renegade" departments from the grade distribution pool, the mean UW-L GPAs would rise even higher than the statistics cited above indicate. In addition to this problem, institutional records also show that in almost all cases where multiple sections of a single course are offered, the grading distributions vary widely from instructor to instructor. This clearly indicates that the grade a student earns in a course can be highly dependent on the section in which he or she has enrolled, and not on the mastery of course content.

Thus, two major problems are evident from this brief analysis: 1) in general, the grades assigned to students on our campus are becoming so highly inflated that they are beginning to lose their meaning and value and 2) the lack of consistency in grading practices, both across and within, departments creates a situation that is unfair to our students. In this paper, we will address these two issues in greater detail and make some recommendations to remedy the current situation.

Before proceeding, we wish to make it perfectly clear what this position paper is *not* attempting to do. We will not address the derivative question of what is driving grade inflation (*i.e.*, the motivating forces behind it). As interesting a question as this may be, it is not necessary to establish the etiology of contemporary college grade inflation to determine whether or not something should be done about it. The question of whether or not college grade inflation merits correction is a separate question from determining what factors have converged at the present time to bring about this phenomenon. To fail to separate these questions would adversely impact our capacity to have a meaningful discussion on either one of them.

### **A Brief History**

If the mean GPAs reported above strike you as excessively high that may be because they are. Back in the 1960s, there was no issue of college grade inflation because it was normative at the time (at least among the more prestigious colleges of the nation) to grade on a bell-shaped curve or some facsimile thereof. This policy had the virtue of assuring that there would be a reasonably broad and regular distribution of college grades, regardless of how brilliant the students in any particular college class might be. Students were simply compared to the others in their class (their “peer group”) in assessing their relative mastery of course content. This policy also assured that the overall mean grade would hover around 2.00, the so-called “gentleman’s **C**.”

Today, students with a much broader spectrum of ability attend college; a fact that has resulted in a drawing down of the national mean SAT scores. In fact, this “deflation” of SAT scores became a big enough concern that the industry that manufactures the SAT tests felt compelled to re-norm the test scores during the prior decade. In short, SAT scores were artificially inflated to raise the national mean SAT scores to be more in accordance with former SAT averages. Paradoxically, these re-normed SAT scores are sometimes cited as justification for the current elevation of student course grades.

Nonetheless, it is a truism that the distribution of college grades 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. In other words, the distribution of **A**, **B**, **C**, **D**, and **F** grades, or any other grading system, is a fundamentally arbitrary matter to be determined entirely at the discretion of course instructors. The traditional rubric for grading has been to use letter grades to describe a student’s mastery of the course material. In this rubric, the letters have the following meanings: **A** = “outstanding, well above average work,” **B** = “good or above average work,” **C** = “satisfactory or average work,” **D** = “barely satisfactory or poor work,” and **F** = “unsatisfactory or below minimal standard work.” The percentages of a particular class’s students meriting these letter grade markers has always been left up to the college instructor to determine. However, 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. Regarding the traditional grading rubric, we have now reached the point where nearly all UW-L students are “above average.” In short, the academic world at our university now approximates the mythical “Lake Wobegon” created by Garrison Keillor, one of our foremost national celebrities, where “all the children are above average.” *Of course, Garrison Keillor is a professional humorist, and his tongue-in-cheek joke is that it is a logical impossibility for all students to be “above average.”*

We have also begun to observe that one unforeseen consequence of escalating grade inflation is a phenomenon that may be termed “*grade rage*.” While space restrictions do not permit extensive elaboration; suffice it to say that “grade rage” refers to the hot-tempered outburst that increasingly now arises when a student receives a lower grade than what the student has grown accustomed to receiving—even when that grade drops as low as an **AB** or a **B**. Such students are known to fly into an irrational attack upon the integrity of the instructor who has dared to give the student the grade that has actually been earned. In our experience, these students are quick to argue that they “have always received **As** or **ABs**,” and therefore the instructor must be in error. Such students never seem to be interested in discussing the demonstrable flaws in their work, but prefer instead to argue “in principle” that all of their work “deserves” an **A** or **AB** grade. If the instructor is honest with such a student about the deficiencies in the disputed work, it becomes readily apparent that this student has internalized a world-view in which lower-end grades are virtually impossible to attain. Such is “grade rage”—a phenomenon not often encountered in the academic world until our contemporary times of grade inflation.

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

Recognizing that college grade inflation is no rhetorical invention, we must now answer our first substantive question; namely, “What exactly is wrong with grade inflation?” Grade inflation seems to resonate well with our popular culture and its “I’m OK, You’re OK” mentality. It also resonates well with our culture of consumerism, wherein certain students and their families believe that they have the right to be awarded high grades as well as a college degree simply because they have paid for their college experience. No doubt you have already discerned that this paper is going to argue that there is something terribly wrong with any college grade distribution policy that fails systemically to make meaningful and clear distinctions between “exceptional,” “above average,” “average,” “below average but satisfactory,” and “unacceptable” levels of performance in college courses. However, in order to present our case for giving “meaningful grades,” we need to address a variety of issues pertinent to the contemporary grade inflation trend.

Superficially viewed, grade inflation appears to be relatively harmless. After all, it does provide more praise and honor for more students. Who would dare to find fault with a grading policy that eliminates what so many college instructors find to be the single most difficult part of their jobs—giving their students “meaningful grades?” In fact, nothing would be more personally gratifying for us than to be able to give all of our students grades of **A**. *The problem is that to do so would not fairly represent the actual student performances exhibited.* Indeed, we have often hoped that the thousands of college students we have taught over the years would all earn **A** grades on their first course exam, but this has never even come close to happening. The reasons for this are manifold, as we are certain any experienced college instructor can appreciate. Not to be ignored among these are the simple truths that some students work harder than others, and some students are just better able to integrate the course materials into their prior accumulation of knowledge. It would be both naïve and absurd to assume that students enter our classrooms *tabula rasa* (a Latin/English phrase signifying a mind unaffected by any prior life experience). And, as all of our professorial experience amply bears out, few, if any, of our students exit our classrooms with exactly the same level of course mastery.

The grades we use to describe a student’s course mastery are merely symbols or messages. They are the practical means that we use to tell students, and subsequently, parents, employers, and graduate and professional schools, how well those students have done in our

courses, relative to their peer group. *Grading that misrepresents student performance sends a false message, and it robs the best students of credit for their good work. It is misleading, and it eventually strips the faculty member and the university of professional credibility.* Inflating grades to please students (and possibly their parents) is not only confusing but is ultimately self-defeating. When an **A** grade no longer distinguishes “outstanding” from “good,” teachers lack a formal means to inspire the passionate striving that can motivate students to achieve their highest academic potential. At the other end of the spectrum, issuing a grade of **F** has now become more of a disciplinary action designed to penalize students for “failing to come to class” or “failing to submit assignments,” rather than for “failing to do acceptable college-level work.” In the final analysis, when the highest grade ceases to identify a student’s work as “outstanding,” and the lowest grades are rarely used, the power of faculty to recognize and reward appropriately is lost. The tragic result for students is that they do not receive accurate information about where they must work more to improve their skills or where they are performing adequately. Kamber and Biggs have recently described this situation as “grade conflation,” rather than grade inflation, due to the fixed upper limit (the **A**) of the traditional grading rubric. (*The Chronicle Review*, April 12, 2002.) And with grade conflation comes a loss in the value and meaning of grades altogether. This forces those who rely on grades to do their jobs well (employers, admissions committees, scholarship review committees) to look to other, perhaps less meaningful, indicators in choosing students. Again, the faculty lose their ability to accurately signal to others which students are most likely to succeed in a given professional area (usually an area in which the teacher is a knowledgeable expert), and the students do not receive fair treatment. We therefore suggest a return to using the more traditional definitions of our “**A** through **F**” grading scale to restore the utility and value to this system.

To summarize, grade inflation does indeed matter. One thesis of this short paper is that it is one of the faculty person’s foremost professional obligations to award course grades that reflect the relative degree of student mastery and achievement of course objectives. It is our considered opinion that if this cannot be done, then there must be something either wrong with the course (e.g., it is not sufficiently challenging) or with the instructor’s competence as it relates to grading ability.

### **Shouldn’t All Students Earn High Grades?**

Now, the question naturally presents itself as to why we must be forced to rule out the hypothetical possibility that all (or nearly all) of the students in a given classroom could not be so inspired by their instructor and their career ambitions that they all earn the highest grades possible (**A** or **AB**), despite their heterogeneous backgrounds and unequal academic preparations prior to entering a particular course. To this question we are prepared to concede that not only might this happen (“in the best of all possible worlds”) but it may be, in many instances, predetermined to happen. There are at least two diametrically contrasting ways that an entire class might all receive grades of **A** or **AB**. (Henceforth, we will refer to these grades with the phrase, “*top end*” grades.)

In the first way, the entire class somehow manages to eradicate all those naturally occurring differences in student aptitude, motivation, class attendance, effort, and accumulated stores of relevant knowledge and insight as these bear upon the quality of the student’s aggregate course performance. (Possible, but highly improbable.) In the second way, the course instructor simply lowers the bar of course expectations to such a level that every, or nearly every, student earns these “*top end* grades.”

If pressed to explain which of these two scenarios might better explain how it happened that some particular class ended up with a more or less uniform “top end” grade *distribution* (all **As** and **ABs**), it is unlikely that any instructor would volunteer to admit to the latter of these scenarios. After all, this scenario is tantamount to an admission of an abdication of professional responsibility because the meaningful grading of student work is a primary responsibility of every college faculty member. In this scenario, course grades become virtually meaningless as symbols conveying relative degrees of course mastery. However, it should be noted that this reality might not be appreciated by any particular person reviewing the grade transcript of someone from one of these classes. This is because it is still assumed by most adult Americans that “top end” grades are scarce and signify outstanding course performance as compared with all other students taking the same course. *And herein lies one of the most grievous wrongs with uniform top end grade distributions: they actually provide no meaningful information about student performance, and they create an erroneous transcript record of “excellence” in course mastery that is tantamount to academic dishonesty* (barring, of course, some notation on the transcript that top-end grades were given to virtually everyone who completed this particular course). Understandably, the proposal to make such a transcript notation would be vehemently opposed by all those who seek to benefit from “top end” grades without having to actually earn them.

It would appear, then, that the only remaining avenue through which to seek a justification for giving uniform “top end” grades is the patently suspicious claim that *all* of the students have in fact achieved “top end” course mastery (scenario 1)—a distinction that in past decades has only been attainable by a small percentage of the student body. What are we to make of such a claim? Let’s consider some of the credibility problems that this claim raises. First, we must recognize that it won’t work to argue that maybe all the best students in the university end up being in the same class by some freak coincidence, and therefore, it is entirely appropriate that these students all receive “top end” grades. This argument won’t work because the grade inflation trend that is being addressed in this paper applies to the *entire university* and not just a few isolated courses or instructors.

We do not deny that the problem of grade inflation is greater in some parts of the university than in others, but the fact remains that the aggregated grading patterns of the *entire* university are what define the problem. And, given that the overall UW-L student body is undeniably heterogeneous with respect to the ACT scores they bring with them, the high school class ranks they have earned, the foreign language proficiencies they exhibit, the number of study hours they invest in courses, and so on, the expectation of some reasonable variance in the grade distributions for these students seems, to us, to be irresistible.

### **Don’t Rising ACT Scores Translate Into Higher Grades?**

Another attempt to defend the overall inflating of grades takes the position that because the median ACT scores for UW-L students have consistently risen over the past decade, grade inflation is an inevitable and justifiable by-product. While it may very well be true that our students are getting better, we argue here that *course performance expectations not only can, but, in our opinion, should rise as the capabilities of the students rise*. Our students may indeed be above average with respect to the general population, or even the general population of students attending college in Wisconsin, but *within their own peer group here at UW-L*, how do they compare? Do the grades we assign accurately reflect the relative degrees of mastery of course content? Again, we feel that it is our responsibility to adequately

challenge students of all ability levels and to then provide honest feedback with meaningful grades that make acute distinctions in the quality of their work.

### **Shouldn't Undergraduates Be Treated Like Graduate Students?**

Another argument that has been raised in defense of a uniform “top end” college grade distribution invokes the analogy of graduate school. The argument made here is that if high grading is not problematic in doctoral graduate schools, why should it be any less problematic for undergraduate students? We concede that graduate school grading policies do indeed tend more towards “top end” course grading distributions. However, we dispute that there is any good foundation for comparing undergraduate grading policies with doctoral school grading policies. Simply put, doctoral graduate schools accord relatively little importance to course grades because these are supplanted by several other measures of progress that include mandatory preliminary examinations, an oral dissertation defense, and a written doctoral dissertation, among others.

Perhaps more importantly, graduate students are the *crème de la crème* of the undergraduate students of this nation, and they have much higher expectations placed on them than do undergraduates. Thus, while course grades are of relative insignificance to a graduate student, these same grades have a much more utilitarian value to the undergraduate student because they are used to determine who will or will not be placed on academic probation, who will or will not be eligible for various academic scholarships, who will or will not likely receive a good recommendation as a prospective graduate student, *et cetera*. In general, undergraduate students are a much more heterogeneous population than are graduate students. This is reflected in the wider range of their educational preparations, their wider range of undergraduate career aspirations, their wider range of academic aptitudes, and, for many, their variable purposes for even being in college in the first place. For these and many other reasons, undergraduate grading policy should reflect this greater range of abilities and seek to provide meaningful information, both to the student and to others, on the relative degree of academic achievement that has been demonstrated.

### **Should Grades Be Raised to Improve Retention?**

Yet another argument in defense of a “top end” course grade distribution is that it promotes elevated college student retention rates. While we think that this point is undeniable, it nonetheless raises the ethical question of whether the end can justify the means. In short, if students who are performing poorly in the classroom are nonetheless receiving “high grades” (or at least grades high enough to avoid academic probation), then they are keeping other students out of the classroom who are more deserving of admission into the university, or at least, deserving of the chance to prove themselves. This is necessarily true given that we continue to have limited enrollment regulations (*i.e.*, enforced enrollment caps). In summary, this argument is ethically tainted.

### **Grading Disparities and Fairness**

Finally, it must be stated that “top end” grade inflation compromises the integrity of the university in a number of additional ways. To the extent that the aforementioned grade inflation is uneven across disciplines, some students (*i.e.*, those inside less grade-inflated disciplines) are placed at a relative disadvantage to other students (*i.e.*, those inside more grade-inflated disciplines) whenever the overall grade point average is used as the basis upon which student scholarships are administered. As noted earlier, grade inflation seriously undermines the distinction of graduating with honors. (“When everyone graduates with

honors, no one does.”) To the extent that grade inflation is uneven across disciplines, there is also an implicit inducement for some academic programs to engage in a higher level of grade inflation than the competitor disciplines in order to attract more students. Similarly, *within* departments where multiple instructors teach different sections of the same course, students do not receive fair treatment if the instructors use highly different criteria for assigning grades. In this case, a student’s grade depends too heavily upon the section of the course selected, and not on the absolute mastery of course content.

In sum, we believe that the many negative consequences of grade inflation require that some action be taken to restore meaning and value to the grades we use. This action should improve our ability to accurately communicate student mastery of course content and should seek to level the academic playing field for all of our students.

---

### **Action Recommendations**

This brings us to the matter of recommended remedies; namely, “What might effectively be done to curb the current levels of college grade inflation as well as the wide disparities in grading that exist across different sections of our courses?” While there are undoubtedly a variety of possible remedies for the first problem (grade inflation), we are convinced that there are two primary solutions readily available to us. If either of these solutions is followed, the second problem (grading disparities) should also be solved. However, if neither of our recommendations is taken, the second problem may still be corrected rather easily.

To address the second problem (wide variations of grade distributions among different course sections) we simply recommend greater levels of communication between instructors. Department chairs should be encouraged to share grade distributions for all departmental courses with the teaching staff. Additionally, instructors teaching the same course should discuss with one another, and agree upon, some common criteria to be used in assessing mastery and assigning grades. Frank, open and honest discussion of these topics could go a long way in leveling the playing field for our students.

We now proceed to addressing the larger problem of grade inflation in general. Before looking at our recommendations, let’s first consider one oft-mentioned solution and explain some of its drawbacks.

**Elimination of Grades.** We have noted earlier that one of the most difficult tasks of a teacher is to assign fair, meaningful grades that communicate the level of course mastery to a student. In general, faculty do not enjoy grading; and likewise, students typically do not like to be graded. So why not eliminate grades altogether? Graduate programs clearly place less emphasis on course grades; why shouldn’t undergraduate programs?

There are a few important reasons why we must not eliminate grading. The first is that we must retain a means to determine if a student is actually deserving of the baccalaureate degree. By communicating individual course mastery to students through grades, a growing record of both student achievement and relative ability in a wide variety of subjects is created. This record then serves as documentation that a student has indeed completed a specified course of study that is both broad and deep. It also indicates the level of student mastery in each subject. To eliminate grades would force us to create other means of assessment that might accomplish the same basic objective. These alternatives have not been forthcoming. Additionally, in the very heterogeneous undergraduate population, grades serve as a highly efficient way to record and communicate student achievement and relative ability in the classroom. Grades are already well established as the accepted “legal tender” of academic performance, and are used to send important signals to others downstream (*i.e.*, employers, admissions committees, *etc.*). While graduate programs have developed other means of tracking student development, they have the benefit of serving a much smaller, less heterogeneous population of students—one that has supposedly already been screened and filtered at the undergraduate level. Eliminating grades may appear to be an attractive, easy fix for the problem, but developing a suitable alternative to the traditional grading system would still remain as an institutional imperative. Indeed, virtually all institutions that have experimented with alternatives have soon returned to the traditional system that has worked quite well for many, many decades.

**Transcript Notations or Increased Use of Pass-Fail Grades.** This approach could involve a variety of simple and inexpensive modifications to our transcript that would clearly convey the type of grade distribution an instructor has chosen to use in a course. For example, the percentage of all students earning a particular letter grade in the class could be reported next to the student’s grade. Alternatively, the calculated mean GPA for the course could be reported in numerical or letter format. Beyond these solutions the use of **pass-fail** grades could be increased. For example, in those courses where top-end grade distributions (all **As** and **ABs**) are exclusively assigned to all students, instructors could be encouraged to use the **pass-fail** rubric instead of the traditional scale. This would leave the full “**A** through **F**” scale to those courses where all of the letter grades are typically assigned to students and where the traditional meanings of these letters is implied. Any of these potential solutions would serve to restore some meaning to the traditional letter grades without asking faculty to severely alter their current grading practices. Transcripts would simply become more informational.

One benefit of this solution is that courses in which relative student ability and student rank *is* considered in assigning grades could be more easily identified. Likewise, one could also identify courses in which ranking students is either impossible or irrelevant. In general, transcript reviewers could find the types of information they seek much more easily, and the traditional letter grades could become useful signals once again. The ultimate value of this proposal, however, is that it is simply a more detailed and informative way of reporting objective information about student performance. Further, it would serve to more effectively reward students who consistently do well in their coursework relative to their peers. It would help us to better recognize those students who learn and exhibit the behaviors we attempt to teach, rather than to deflate their worthy accomplishments. Finally, this solution would eliminate the deception inherent in situations where awarding a “top end” grade to a student signifies that the student has done “*outstanding, well above average*” work, but in fact the “*average*” course grade is also the “top end” grade awarded. These situations ultimately do a

disservice to our students, and it is hoped that they, and any other transcript viewers, would appreciate receiving information that is a bit more accurate.

**Establish Grade Distribution Guidelines.** This final potential solution could target either all UW-L undergraduate courses or only UW-L General Education courses. It would stipulate that grade distributions must be centered around a mean grade of **B** (3.00) or lower. Exceptions to this policy option could be provided, but exceptions should be relatively few in number and should have to be justified to the Dean of the College in which the course is housed. If this simple remedy were to be adopted, it would be entirely up to the course instructors to make grading policy adjustments, if any were needed. As envisioned here, there would be no binding compulsion that the recommended grading distribution guideline be precisely conformed to every single time a course is taught. However, there would be a forceful expectation that, on average at least, instructors would closely approximate this recommended grading guideline (*i.e.*, norming course mean grades at 3.00 or below). For example, if an instructor wishes to award a grade of **B** (3.00) to 90% of a class, then roughly 5% **As** (4.00) and 5% **Cs** (2.00) should be assigned to the remaining students.

It would be expected that failure to conform to this grading policy guideline be addressed in department merit evaluation procedures, by oversight of the college Dean, and by promotion and retention considerations at the university committee level. Hopefully, voluntary compliance with this recommended grading policy would render the need to use any sanctions completely unnecessary. However, 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. It is in this spirit only that the aforementioned sanctions are stipulated.

---

In conclusion, we believe that we can better promote the university's mission of educational excellence, if we were to take some form of corrective action as described above. We should not lose sight of the fact that it is possible to remain merciful and benevolent in our grading without abdicating our professional responsibility to our students, to our colleges, or to our university. In addition, those people who employ us and pay our salaries, and those who use the information we provide about our students expect nothing less from us as teachers.