

Grades and Grading in Higher Education and at UW-La Crosse: A Discussion Paper
by Robert Bilby, Department of Sociology/Archaeology

Some faculty members and administrators have expressed concern over the distribution of grades at UW-La Crosse, and in higher education more generally. Our answers to three key questions largely determine whether we see the results of our grading practices as a problem.

- a) Has a trend toward assigning higher and higher grades in recent decades reached alarming levels?
- b) Have we at UW-La Crosse been part of such a national trend by becoming increasingly easy graders relative to our students' academic skills?
- c) More generally and probably most importantly, what do grades mean and how do they function in the learning process?

This discussion paper summarizes research findings that provide what may be surprising answers to each of these questions in turn. [In a conscious effort to not bog down in technicalities, endnotes are used for the more detailed or technical information judged to be less central to the main points.]

National Trends and Perceived National Trends

The belief that university and college instructors now give far more high grades and far fewer low grades than was the case two or three decades ago seems to have achieved the status of verified truth. The types of evidence that have been used by the most vocal national critics, however, are quantitative studies of single institutions or a few institutions – usually elite universities – or else are anecdotal observations.

But just last summer, the U.S. Office of Education published a study finding that the distribution of grades given at four-year public universities in academic year 1999-2000 are not as problematic as the critics feared. This “Profile of U.S. Postsecondary Education Institutions: 1999-2000” (<http://www.nces.ed.gov/pub2002/2002168.pdf> ; or see the *Chronicle of Higher Education*, July 12, 2002:A37) reports data, shown below, from the 1999-2000 transcripts from a representative sample of about 62,000 students attending more than 900 post-secondary institutions.

UNIVERSITY GRADE-POINT AVERAGES NATIONALLY, 1999-2000
(U.S. Office of Education, 2002)

| | C's and D's or lower | B's and C's | Mostly B's | A's and B's | Mostly A's |
|---------------------------------|-----------------------------|--------------------|-------------------|--------------------|-------------------|
| ALL UNIVERSITIES | 33.5% | 16.4% | 24.6% | 10.9% | 14.5% |
| SCHOOL TYPES | | | | | |
| Public 4-year | 34.4% | 21.2% | 25.1% | 9.7% | 9.7% |
| Public Non-Profit 4-year | 22.4 | 18.2 | 30.1 | 14.4 | 15.0 |
| Public 2-year | 38.2 | 13.2 | 22.3 | 9.8 | 16.6 |
| All-Private for-Profit | 25.6 | 12.1 | 23.8 | 16.2 | 22.3 |

We can see that the total of “mostly A’s and B’s” and “mostly B’s” is about one-quarter of all grades given at all types of schools, and less than one-fifth of grades given at 4-year public schools. “C’s and D’s or lower” are given one-third of the time at all types of schools, and slightly more often at 4-year

public schools. The grade distribution for public 4-year institutions has the lowest proportion of high grades. While some people still may be troubled by these finding on grade distributions, it seems that, as noted in the Chronicle, “despite allegations of rampant grade inflation, the C is alive and well.”

But how could this be, in light of the sincere and heartfelt criticisms of grading practices that have so captured the climate of opinion in recent years? Apparently the critics are focusing on a problem that exists primarily in a small segment of higher education. Findings from another U.S. Office of Education study published in 1995 (see Adelman, 2001), this one a longitudinal study of a large sample of college transcripts, shows that nationally both grades and grade-point averages did not increase between 1972 and 1993. The inconsistency between this fact and the perceptions of the those so alarmed about grades is due to overgeneralizations from the situation at Ivy League and some other elite schools where instructors in fact give very high proportions of their students high grades.

More specifically, among students who received bachelor’s degrees over the period of 1972-1993, the average GPA actually declined from 2.98 to 2.89.¹ Elite schools were giving increasingly more high grades over this period, but that does little to skew the overall numbers because less than 2% of all undergraduates and less than 5% of all bachelor’s degree recipients attend elite universities.²

So the reason that research on the grading situation nationally in 1999-2000 does not show the result of a ubiquitous trend toward ever-higher grades is that such a trend never happened! The lesson here for us is that the discussion of our grading practices should not be framed by a heart-felt concern over too many high grades being given at Harvard and some similar schools, coupled with the misperception that it’s that way almost everywhere else.

However, other than the comparisons among broad types of universities, these studies of average grades given by large numbers of colleges and universities do not systematically address the question of how much variation in grading occurs across institutions. We do know that at UW-La Crosse, average grades have wavered upward over the years. Let’s turn to the situation locally.

The Grades We Give

Data on grades given at UW-La Crosse from Fall, 1975 through Spring, 2001 were presented last year at a forum on grading held at the beginning of the Fall 2002 semester.³ These data show that there has been an increase in the average grade given over these years.

To establish precisely what we are discussing, let’s look at the magnitude of the change in the all-university average grades over these years. But first, because average grades given in the fall semesters are always lower than grades given in the spring semester of the same academic year,⁴ it is important to compare grades given in fall semesters with fall semesters, and spring semesters with spring semesters. Thus the trend over the academic years we are examining are

Fall, 1975 to Fall, 2000: 2.84 to 3.12, or .28 of a letter grade
 Spring, 1976 to Spring, 2001: 2.86 to 3.19, or .33 of a letter grade

We are looking at a change in the average grades given at UW-L of around three-tenths of a letter grade over a 26-year period.⁵

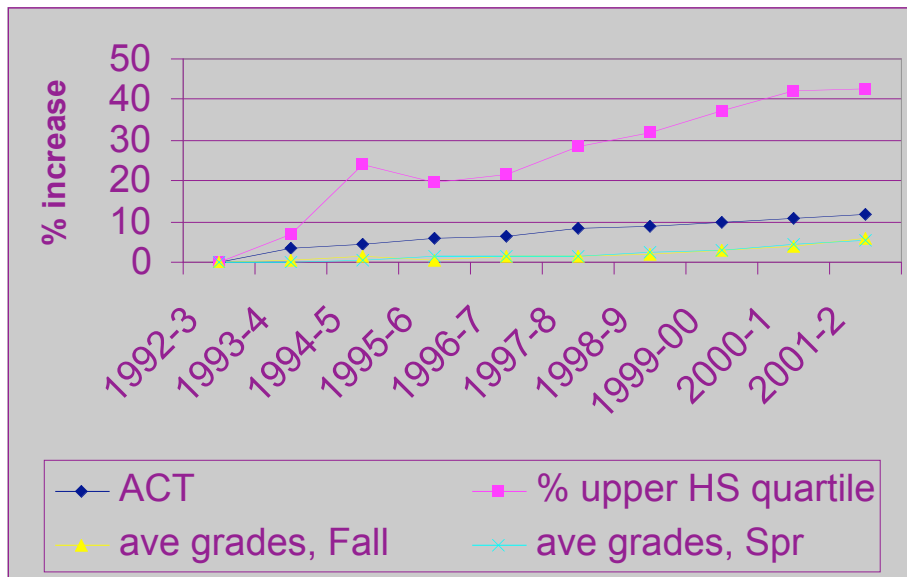
Those who are alarmed by this trend interpret it as reflective of declining educational standards. Yet the changes in the grades that have been given in recent years might well be a function of the

changing levels of academic skills among our students. Each academic year, spokespersons for our University have proudly and truthfully announced the ever-increasing levels of the academic qualifications of our incoming students. Average ACT scores have gone up almost every year for some time now, as have the proportion of our first-year students who graduate in the top twenty-five percent of their high school classes. Let us not discount the impact of this pattern or minimize its relevance without examining it more closely. Indeed, this is an empirical question more than a point of conjecture. Here is one way of looking at the data since 1992.⁶

| | 1992-3 | 2000-01 | per cent change |
|---|--------|---------|-----------------|
| All-university average grades | | | |
| Fall Semesters | 3.04 | 3.12 | 2.6% |
| Spring Semesters | 3.08 | 3.19 | 3.5% |
| Average ACT scores of beginning students | 21.7 | 24.4 | 12.4% |
| Class rank of beginning students (% in top one-quarter of HS class) | 47% | 70% | 48.9% |

This same information is presented in the graph below, including the data on the years intervening between 1992-93 and 2000-01.

INDICATORS OF ACADEMIC PREPARATION AND ALL-UNIV AVERAGE GRADES



Clearly, the quality of the academic track record of our incoming students has increased considerably more than have the all-university average grades these students receive after they enroll at UW-L. While the grades we have been giving over these years indeed increased somewhat, indicators of the grade-earning ability of our students have increased considerably more. One could even interpret this to mean that students are being graded harder relative to their academic skills than was the case a decade ago.

In the discussions on campus last year, the above explanation of grading trends at UW-L was mentioned (though the specific numbers were not consulted). In response, some took the position that we should not place much store in these relationships between, on the one hand, higher admissions

standards and increasing academic skills of our students and, on the other hand, trends in grade distributions. In this view, if our students really are better, we should “raise the bar” in response, and use more stringent grading strategies to keep the grade distribution from increasing. Such a response is seen as part of not letting educational standards deteriorate. This position seems odd in comparison to a logically similar concern on campus. At UW-L, since some programs and colleges give higher average grades than do others, and since for our students qualifying for scholarships is to a significant degree dependent on GPA, the differential grading pattern by program or college results in a disadvantage for those students in the programs that give fewer high grades. This is indeed a thorny problem, for we do not want any of our students at a structural disadvantage. At the same time, then, why should academically more competent UW-L students be given lower grades relative to their ability than are students at other UW-System universities, since our UW-L graduates will be competing with graduates elsewhere for various post-graduate opportunities that sometimes depend on cumulative GPAs? To anticipate later discussion, educational standards involve far more than the grades we give.

Summing up so far: UW-La Crosse is not caught up in an irrational national trend toward giving higher grades because there is no such national trend; there is only a sensationalized phenomenon at highly publicized schools where huge percentages of students do get high grades. Average grades at UWL have gone up about .3 of one letter grade since 1976 and about .1 of one letter grade since 1992. The most plausible explanation, based on evidence about our students, is that our students are better at earning grades than they used to be. What this means to us depends on our assumptions about the meaning and function of grades, to which this discussion paper now turns.

The Meanings and Functions of Grades

The discussions about grade distributions that occurred last year revealed substantial differences among instructors, and among students, in assumptions about grades. A dialogue on grades would seem of tremendous value to the University community since it leads us to consider the very core of our shared purposes as an institution of higher learning – the teaching and learning process, and the relationship of the university to external audiences interested in our students. This involves, at a minimum, examining our assumptions about 1) grades as rewards and/or incentives, 2) grades as feedback to learners, 3) grades as communications to prospective employers and graduate schools, and, finally 4) grades as indicators of educational standards.

Grades as rewards and/or Incentives: Do grades reward learning?

Presumably, most educators take for granted that students experience good grades as rewards, and that grades serve as incentives for our students to learn. Behavioristic psychology has been popularized along these lines. Common sense and personal experience suggest that people work for and respond to rewards. Yet this is a gross oversimplification.

A body of published research going back to the early 1960’s and continuing to the present, using experimental designs with subjects of various ages, shows that people who are offered rewards for various types of intellectual or creative effort do more poorly (make more mistakes, learn less, produce work judged as less creative, etc.) than subjects who are not rewarded (see Kohn, 1993: 35-95). One psychologist’s summary of numerous such published reports states that when offered rewards as enticements to behave, people

choose easier tasks, are less efficient in using the information available to solve novel problems, and tend to be answer oriented and more illogical in their problem-solving strategies. They seem to work harder and produce more activity, but the activity is of a lower quality, contains more errors, and is more stereotyped and less creative than the work of comparable nonrewarded subjects working on the same problems (Condry, in Kohn, 1993).

At one level, this seems preposterous. We know that candy, toys, privileges, grades, salaries and pay incentives, job security and many other types of rewards all influence people's behavior. Many of our society's educational and management practices are based on the routine use of rewards. Yet the precision with which the above quotation echoes the way so many high school and post-secondary students are described by their frustrated instructors is remarkable. *We need not argue over whether rewards make a difference, for clearly they do; but we do need to examine much more closely the nature and effectiveness of rewards, or more specifically, grades.*

Rewards deflect attention from the activity being rewarded (for our purposes, learning) to the reward itself (i.e., grades). A great deal of evidence from experiments has accumulated – most recently a detailed review of 128 such studies⁷ – that shows when activities that are done for their own sake are rewarded, the original “intrinsic motivation” erodes. Rewards, especially verbal praising, function best when they are unexpected, and are provided in a non-controlling manner (i.e., seem spontaneous, sincere and natural). If external rewards are experienced as part of a system of control, they *decrease the person's sense of self-determination*. When that sense of self-direction erodes, even finding out that we have been judged as competent comes to be experienced as game-like, superficial, or even potentially threatening.

Three similar and inter-related concerns about grades should be mentioned (see Kohn, 1993: 54-67). One, the emphasis on grades leads to a focus on outcomes, products, the one right answer, and a de-emphasis on reasons for the answer or the process of creating outcomes. Two, grades discourage risk-taking. And three, grades maximize the power-related aspects of the teacher-student relationship, and push student-student relationships toward either isolated competition or co-conspirator for grade-earning. Indeed, when rewards are perceived to be embedded in a system of social control, they engender a) avoidance; b) strategic or manipulative responses; or c) revolt.

A very similar analysis is based on a distinction between grade-earning behavior and learning. Good grades reward excellent grade-earning behavior, but if grade-earning behavior does not overlap extensively with genuine learning, grades are not rewarding learning. For example, attending classes can be distinguished from engaging in what goes on during those classes; memorizing information and recognizing correct answers on exams can be distinguished from comprehending knowledge and retaining it as part of one's personal cognitive resources. Expending effort to get good grades involves investing energy into figuring out what might get graded and what might not, and then more or less ignoring whatever is thought to fall into the latter category. Certainly some of our students have maintained some degree of intrinsic motivation to learn in spite of experiencing many years of schooling shot through with grading. Yet the proverbial question “Will this be on the test?” and many of the tactical choices students make in striving for grades – including, at times, academic dishonesty – indicate the degree to which authentic learning has been deflected toward grade-earning per se.

The University has spelled out our commitment to enhancing our students' intellectual development, to their critical thinking and problem solving abilities, and to their becoming life long learners. These goals reverberate with the notions of competence and self-determination. Ironically the extreme emphasis that our educational system places on grades as incentives to learn so often

actually gets in the way of the types of learning we associate with “intellectual development,” “critical thinking,” and “a love of learning.” We need to consider seriously the degree to which grade-earning and learning overlap in our classes, and whether extreme emphasis on grades has literally gotten in the way of the University’s goals of fostering various types of intellectual growth.

Do grades give feedback that fosters student learning?

In grading, we evaluate a student’s work or performance relative to some criteria and/or in comparison to the performance of others. Letter grades or numerical scores are general judgments of performance. The grade itself does not contain information about just what a student did well to merit a good grade or did poorly or not at all to merit a low grade. To suggest that grades provide feedback to learners overlooks that the grade does not contain specific information that students can use to improve their performance. The bad grade’s obvious implication that a learner needs to “do better” or “work harder” often is too vague to help the learner.

Indeed, instructors often lament when students simply look at a grade and then either celebrate or feel disappointed, while ignoring the extensive written feedback the instructor invested so much time to provide. This is exactly the response that develops when obtaining grades take on more significance in the student’s life than learning, and is exactly a response that is consistent with the view of grades and grading set forth in the previous section of this paper.

Do Grades Give Useful Information to Employers and Graduate Schools about Applying Graduates?

Prospective employers and graduate program admissions committees use grade-point averages, grades within the student’s major, and other content of transcripts in various ways to make hiring and admissions decisions. Yet how this information is used varies considerably. Some employers rank grades far below other information about graduates. Some use the GPA as a general indicator of how well the graduate has successfully dealt with the challenges and stressors of college. Others use it as a screening device in early stages of evaluating an applicant pool, and pay little more attention to it. Graduate schools, too, vary by discipline and from school to school in the priority that is placed on undergraduate grades.

There is reason to doubt the wisdom of these practices, even beyond the critical analysis presented earlier. While recent research on the predictive power of college grades is difficult to find, studies done in the 1970’s and earlier found that grades predict future grades moderately well, but do not correlate with much else about a person a few years after they leave school. Grades in professional schools are moderately correlated with ratings of occupational success in the early stages of careers. But such expected correlates as occupational prestige, career success in general, and income were not found to be correlated, or were correlated only very weakly, with grades. The world outside of academia places a much greater emphasis on credentials – the degree a person earns – than on academic success while obtaining that degree. While these relationships may have strengthened in recent years, it is not clear why that might be the case.

A broader question implicit in the provision of grades to prospective employers for their use in hiring decisions concerns the use of the educational system as a sorting mechanism, and the impact that use has on the very nature of the university. To the extent that grading is a problematic educational practice, or one that is very seriously overemphasized, the use of grades in the occupational sorting process makes things that much worse. And, if grades are not actually valid indicators of what employers are looking for, they subvert the employer’s purposes as well. In a social

system that places so much emphasis on individual responsibility in so many of our social arrangements, perhaps entrepreneurs and corporations alike are better served if they are made to gather their own information about applicants rather than relying on a rather distant educational system to inform them of who might best serve their purposes.

Grades and Educational Standards

Grades are our evaluations of the academic work our students' complete, or in some cases their performances in a laboratory, on stage, or in a studio. Discussions of educational standards sometimes refer to educators' grading practices, and at other times to the expectations educators communicate to students. Upon closer examination, grades and expectations are not at all the same phenomena. Expectations are communications, messages about what the learner can and should accomplish in the learning process. These can include messages about the knowledge and/or skills to be learned, processes involved in the learning, available resources, and the level of effort the learner can anticipate needing to invest in the process. Grades are after-the-fact judgments of the extent to which the learner has met those expectations.

A more accurate understanding of standards recognizes that they are reflected in the content of the expectations and in the clarity of the messages involved, not in the distribution of grades assigned at the end of the learners' work and effort. Imagine a situation where, after students and teacher have completed a course and final grades have been given, all the grades are systematically lowered. This would not raise the educational standards – it just lowers the grades. While the example is absurd as something that might happen, its point is that grades per se are not the bearers of standards, they are the end result of a complex teaching and learning process. Lowering them will not raise the quality of the process.

A huge research literature on learning climates in classrooms and schools shows that the quality of the learning environment is characterized primarily by

- appropriately challenging expectations, clearly communicated
- consistent, clear, and supportive and rules defining how the learning environment operates, along with appropriate specific feedback to students, and
- students' perceptions that their teachers think they can learn.

Grading needs to be understood in light of not only the pitfalls surrounding grades, but also in the context of these features of positive learning environments.

Recommendations

1. Refrain from any policy changes regarding grades until much more dialogue has occurred and some semblance of a consensus on such policies emerge.
2. Continue the dialogue on grades in a non-adversarial way, seeking creative alternatives to the imposition of grading practices while clarifying the nature of “appropriately challenging expectations.”
3. Seriously explore fundamental alternatives to the traditional system of grading.
4. While taking no steps whatsoever toward vocationalizing the University, consider ways in which the University can better communicate with employers regarding the information they desire about graduates and prepare them for not receiving traditional grade information from UW-L.

ENDNOTES

¹ Clifford Adelman, author of this study, is a senior researcher at the National Institute on Post-Secondary Education within the U.S. Office of Education. He attributes this pattern of virtually no change or a slight decrease in grades over the period of his data to the considerable expansion of higher education. With the number of students pursuing higher education increasing nearly 40% from the early 1970's to the mid-1990's, types of students who had not attended college in earlier periods were now in college and university classrooms. Adelman and others argue that a many of these students did not have the same grade-earning skills as the types of students who had gone to college in earlier decades.

² Social scientists have documented cases where, in circumstances where the actual amount of crime is stable or decreasing, media coverage of worrisome or sensational crimes have created the shared sense that a crime wave exists. In fact, until very recently one would have had a difficult time learning from the six o'clock news or the headlines of the print media that violent crime in the U. S. has decreased considerably for several years, with decreases in most categories of violent crime going down each year since 1994.

³ Tom Hench (Business Management), Donna Anderson (Economics) and Aaron Monte (Chemistry) conducted this forum and made presentations. Our collective thanks go to Aaron Monte for his work in assembling this data on grades.

⁴ This pattern is no doubt due to the absence in the Spring Semesters of students who did not do very well the previous Fall Semester, and who either choose not to return or become academically ineligible.

⁵ While these differences are only slightly smaller than the .35 difference when comparing Fall, 1975 with Spring, 2001, small differences are proportionally more important when we are dealing with a small range of scores – i.e., between 0 (F's) and 4.0 (A's).

Two other related “technical” points also are important to consider when presenting this data, if only because how these average grades are portrayed visually can have a major impact on our interpretations -- especially our first impressions. *These are offered as cautionary, rather than definitive, points.* First, making the appropriate comparisons (i.e., fall with fall, spring with spring) graphically is relevant, since when superimposing lines of best fit on data points on a graph, a relatively lower starting point and relatively higher last point will make the line steeper. With the first point a Fall average GPA and the last point a Spring GPA, the linear relationship is altered to a slightly steeper line. Ignoring even such subtle distortions like systematic Fall-Spring differences in GPA can give the visual impression of a somewhat greater relative change. Technically, a more accurate picture would emerge if we had two separate line graphs, one with only fall grades and the other with only spring grades.

Second, the basis for the curve on the graph of grades given over this period is not clear. The data points on the graph representing each semester's average grades show that there has not been a uniform pace of change in the increase in average grades given over the years, with a decline in the 1980's, a more accelerated increase in the early 1990's, another brief decline in the mid 1990's, and a relatively slower increase in the later 1990's. While of course the benefit of superimposing a line on an array of data points is to visualize the general trend in the data, it is not clear that a curvilinear relationship is more appropriate than a linear one. It is interesting that in the graph titled “UW-L All-University Grades (1976-2001)” the “log function curve” accelerates upward especially steeply by around 1998 or 1999. In the graph used in last year's forum titled “UW-L All-University Grades (The Future),” in which the same data is presented but the time axis is extended to the year 2025, the same pronounced upward curve does not appear on the graph until around 2015. There seems to be an element of arbitrariness here, imposed by the graphing procedure. More simply stated, *there is no reason to believe that the rate of increase in grades we give will increase as the curvilinear relationship plotted on the graph so strongly suggests.* Using a graph in this way can serve unintentionally as a persuasive or rhetorical device as much a neutral description of data.

⁶ Academic year 1992-93 is used here as a starting point because it represents approximately the point at which UW-L made some major transitions in its identity and relative emphases. The allied health initiative was becoming a reality and the college structure of the university was being fundamentally altered. In short, factors that might be still operative today, rather than in the more distant past, seem to have been in place by around this time. However, extending the following analysis into earlier periods might be worthwhile.

⁷ See Deci, et. al. (2001). These did a review and meta-analysis of the large experimental literature on whether “extrinsic rewards” undermine “intrinsic motivation,” the type of motivation reflected in activities people choose to do and in which they find or develop interest. The researchers found that verbal rewards do not enhance motivation if they are used in a controlling way, but do so when provided in an “informational” way rather than with a “controlling interpersonal style,” *especially in studies where college students were the subjects in the experiments*. Three different types of tangible reward conditions (tokens, money, or symbols like grades) were studied; these are rewards contingent upon 1) engaging in an activity, 2) completing an activity, and 3) performing up to a specified standard. All three types of tangible reward conditions – all of which reflect some aspect of what we ask college students to do for grades – were found to undermine motivation. Motivation was measured in two ways, by how long subjects engaged in the activity and by self-reported interest in the activity. (Deci and his associates of course have their critics (see Cameron, 2001), who argue that the undermining effects of rewards are more limited than Deci and his associates claim, but who do not categorically deny the problems of using rewards.)

WORKS CITED

- Adelman, Clifford. 2001. “Putting on the Glitz: How Tales from a Few Elite Institutions Form America’s Impressions about Higher Education,” *Connection: New England’s Journal of Higher Education and Economic Development* 15:3 (January) 24-26.
- Cameron, Judy. 2001. “Negative Effects of Reward on Intrinsic Motivation – A Limited Phenomenon: Comment on Deci, Koestner, and Ryan (2001),” *Review of Educational Research* 71:1 (Spring, 2001) 43-51.
- Deci, Edward, Richard Koestner, and Richard Ryan. 2001. “Extrinsic Rewards and Intrinsic Motivation in Education,” *Review of Educational Research* 71:1 (Spring, 2001) 1-42.
- Kohn, Alfie. 1993. *Punished by Rewards*. New York: Houghton Mifflin
- Shoichet, Catherine. 2002. “Reports of Grade Inflation May be Inflated, Study Finds.” *Chronicle of Higher Education* (July 12) A37.
- U.S. Office of Education. 2002. *Profile of U.S. Postsecondary Education Institutions: 1999-2000*. <http://www.nces.ed.gov/pub2002/2002168.pdf>