

■ William Potter, Melvin A. Nyman, and Karen S. Klumpp

Be Careful What You Wish For: Analysis of Grading Trends at a Small Liberal Arts College, Grade Inflation or Progress?

Abstract

Grade Inflation has been the subject of hallway and coffee room conversations among faculty throughout the nation at various intervals during the past 40 years. Numerous articles have been published attempting to assess both the extent and the causes of this phenomenon. The following case study, conducted at one small liberal arts college, examines grading trends and considers factors that may have contributed to a general increase in various grade indices since 1969. A primary conclusion of the study is that higher grades at the college level present a dilemma to faculty who constantly strive to find ways to help students achieve at higher levels, but who become concerned once there is measurable evidence that progress may have been made. Secondly, the authors hope that the example in this case study will provide a framework for a more meaningful consideration of this issue on other campuses.

The subject of grade inflation has been the source of heated debate on college and university campuses across the country as well as in the national media. These debates bemoan the upward spiral of undergraduate grade point averages, the likely causes of the phenomenon, and the problems that are alleged to result from that spiral. In spite of the voluminous rhetoric about this issue, only a handful of studies carefully address grading practices and trends. (Potter 1979; Bejar and Blew 1981; Kolevzon 1981; Millman *et al.* 1983; and Nelson and Lynch 1984 are examples.)

This issue has been the stuff of hallway and coffee room conversations among faculty at one small liberal arts college as well. Grades were higher in the 1990s than they had been in the 1980s, the 1970s, or the 1960s and speculation about both causes and cures ran rampant. Many faculty simply attributed the higher grades to the impact of student evaluations on faculty behavior. In the interest of supporting an informed discussion about this complex issue, three staff members took it upon themselves to analyze the data that were available. Of special interest here was the intent to compare overall grading trends to the trend for MTH 121, Calculus I, over a 30-year period. The study began with the requisite literature review, collection of data at the local site, analysis of those data, and development of hypotheses to explain the data trends. Perhaps this approach might help inform similar discussions at other campuses.

Grade Inflation Defined

Over the years, a number of researchers have devised definitions for the problem of grade inflation. Louis Goldman (1985) defined grade inflation as the term used to describe an upward shift in the grade point average over an extended period of time. Birnbaum (as cited in Kolevzon 1981) viewed grade inflation as

rising grade point averages indicating that it was easier to get good grades. Millman *et al.* (1983) claimed grade inflation was the lowering in value of As and Bs because of their more frequent use. Despite the variations on the theme represented by the aforementioned definitions, these writers agreed with Summerville *et al.* (1990) who argued that grade inflation cannot be properly defined merely as rising grades. Rather, borrowing from Bejar and Blew (1981), grade inflation might be considered as "an increase in grade point average without a concomitant increase in achievement." In their study, Bejar and Blew showed that between 1964 and 1974, mean grades for men and women in the first year of college increased steadily from about 2.0 to 2.6 while SAT scores declined. They concluded that grade inflation had, in fact, occurred because grade point average increased without a corresponding increase in the SAT.

Dr. William Potter is currently Dean of University College at Ferris State University in Michigan. His research interests include grading trends, student retention, and higher education middle management issues. His Ph.D. in higher education administration was earned at Michigan State University.

Dr. Melvin A. Nyman is Professor of Mathematics and Computer Science at Alma College in Michigan where he has been on the faculty for 20 years. His primary interests are applying mathematical modeling techniques to problems in biology and social sciences, and the teaching and learning of undergraduate mathematics and statistics. His Ph.D. was earned at Michigan State University.

Karen Klumpp is the Assistant Provost and Registrar at Alma College in Michigan. Among her professional interests are the issues of retention and technology-prompted organizational change. She is a doctoral student in higher education administration at Michigan State University.

Bejar and Blew also observed that grade inflation was not a recent phenomenon, having been in effect since 1964, and that the rate of inflation seemed to have diminished between 1974 and 1978 (Bejar and Blew 1981).

Methodology

The Bejar and Blew approach seemed to be a sensible starting point, and so we decided to compare the grade performance of our students to the ACT-Composite and high school grade point average performance by class cohort. Data about first-year entering classes were requested from the admissions office and from the registrar's office. Data about student grade performance were collected from reports already prepared by the registrar. Data about student grade performance in MTH 121 were collected from grade lists on file in the registrar's office. ACT national data were obtained from corporate headquarters in Iowa City and from the company Web site. These data (30 years of it) were loaded into spreadsheet programs and reports were prepared for analysis.

Results: The Quality of Incoming Classes

As is the case at most colleges and universities, applicants to this college are considered for admission on the basis of high school grade point average (HSGPA) and performance on a nationally normed entrance examination. The HSGPA is recalculated to include only academic courses in the areas of English, mathematics, social studies, science, and foreign language. For most of the period under study, moreover, the ACT was the entrance exam taken by the vast majority of each entering class. For early cohorts, however, the SAT was the examination of record and those scores were converted to an ACT-Composite using a table of concordance.

As shown in Figure 1, the mean recomputed HSGPA of first-year students increased by half a grade point from 1968 to 1998. Although the increase in mean HSGPA may indicate an improvement in the quality of later classes, it was decided to look at other indicators as well.

In the early 1980s, the college introduced two changes that were intended to attract better students. First, the faculty tightened admission standards to minimize the enrollment of "high risk" students. Second, the admissions office changed its financial aid strategy to include awards for merit in an effort to attract more students with high HSGPAs and high ACTs. When other qualitative aspects of each cohort were considered, particularly the dramatic shift upward in the ratio of honors students and the downward movement in the ratio of high risk students, it seemed plausible to assume that any improvement in undergraduate grade point performance could be attributed to the enrollment of better students. However, we still had not considered the one normed indicator available to us, the ACT.

Figure 2 shows the trend for ACT-Composite scores by cohort and seems to indicate a qualitative improvement in the last ten entering classes, following a long period of stability during which scores hovered around a mean of 23.5. Of course, it was essential to account for the change in the ACT scoring standard, effective with the Class of 1995 (entering in 1991).

Even after making a one-point adjustment to the mean for each cohort after 1990, it appeared that some marginal level of improvement had been achieved. Nonetheless, it seemed prudent to consider the ACT scores in one other way before accepting this as a benchmark. Using the data provided by ACT about the trend for national scores, we subtracted the national (college bound) mean ACT-Composite from the college mean ACT-Composite for each year of the study to try to determine

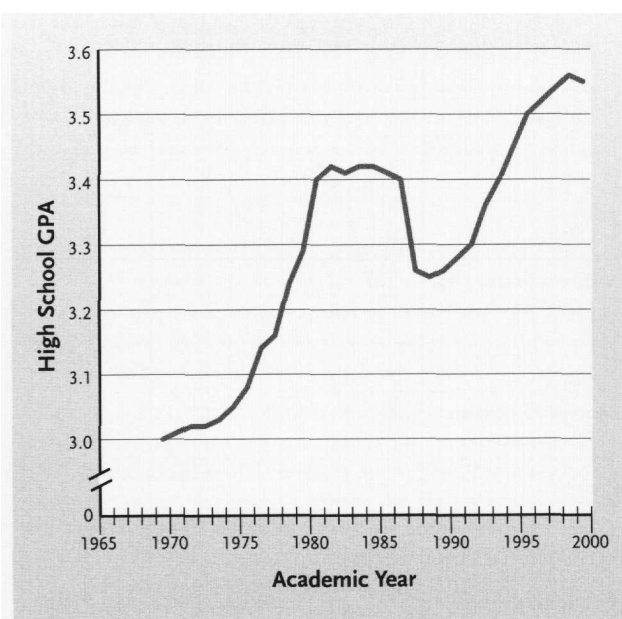


FIGURE 1: HIGH SCHOOL GRADE POINT AVERAGE OF ENTERING STUDENTS, BY ACADEMIC YEAR

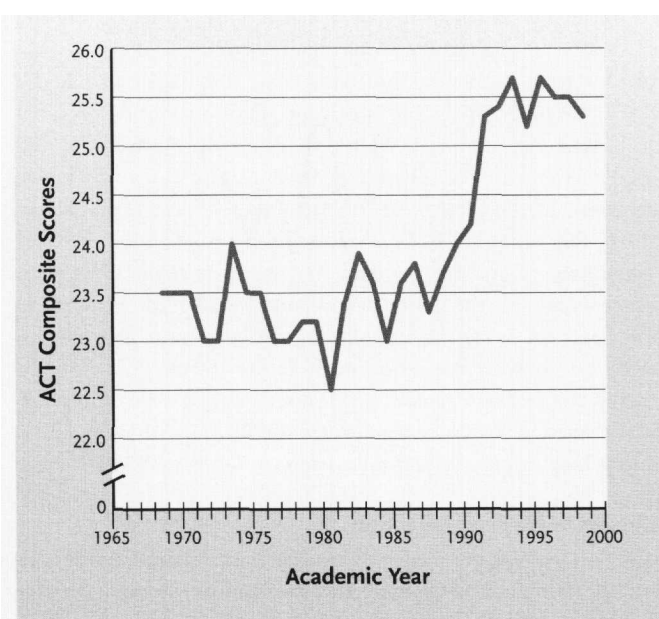


FIGURE 2: MEAN ACT-COMPOSITE SCORE FOR ENTERING STUDENTS, BY ACADEMIC YEAR

whether or not our students had improved. The results of that analysis, shown in Figure 3, were startling. In essence, although there had been some fluctuation of the differences (between three and five points), the difference in the 1970s was about four points and the difference in the 1990s was about four points.

So, although the HSGPA for the entering cohorts had demonstrated a steady and significant upward trend (half a grade) during the past 30 years, careful analysis of the ACT data showed essentially no change, suggesting that the condition to demonstrate that grade inflation had occurred (no concomitant increase in achievement) had been met. Recognizing that both HSGPA and ACT mean are, at best, substitutes for achievement or achievement potential, they were nevertheless the best indicators available for this study and gave us a basis for examination of the phenomenon.

Results: Grade Indicators

The earliest news received about how new students fare academically are the grades posted at the end of the first term. Those data were available since fall 1969 and showed the trend for each cohort without the leveling effects of upper-class students or time in college. (See Figure 4.) The mean first-term GPA in the early 1970s was at or about 2.5. During the 1980s, this indicator increased and plateaued in the 2.75 range. During the past ten years, however, the mean first-term GPA approached 3.0 and actually reached that mark in 1995. Thus, the trend line for what must be considered the most vulnerable and visible stage of development for these cohorts indicates a propensity toward higher grades.

In order to consider how the entire student body had been affected, we decided to review the proportion of high grades

(“A”s, “AB”s and “B”s) given year-by-year. As shown in Figure 5 on the following page, since 1972, the proportion of high grades increased by 20 percentage points.

Students benefited from this largesse with the proportion of students qualifying for inclusion on the Dean’s List (3.50 or better) tripling since 1969.

We also decided to look at the longer-term impact of higher grades by examining the grade performance of each cohort as a graduating class. As shown in Figure 6 on the following page, the median GPA increased by approximately half a grade during the 30-year period under study.

The most striking result from this analysis had to do with the comparison of mean and median cohort GPA increases during the period under review. In particular, between 1969 and 1999, the *mean HSGPA* increased by half a grade, the *mean first-term GPA* increased by half a grade, and the *median graduate GPA* increased by half a grade.

Results: But What About Calculus?

Because one of the researchers is a professor of mathematics, the research team thought it both prudent and interesting to use students enrolled in MTH 121, Calculus, as a way to consider the issue of grade inflation. This first term calculus course for students of mathematics, science, and engineering is a standard course on most college and university campuses. The mathematical content and level of rigor in this course does not vary significantly among mathematics departments in North America. Moreover, according to the chairperson of the mathematics department, the content and level of rigor have not changed much locally over the past 30 years. Although the mathematics faculty had changed completely in the period

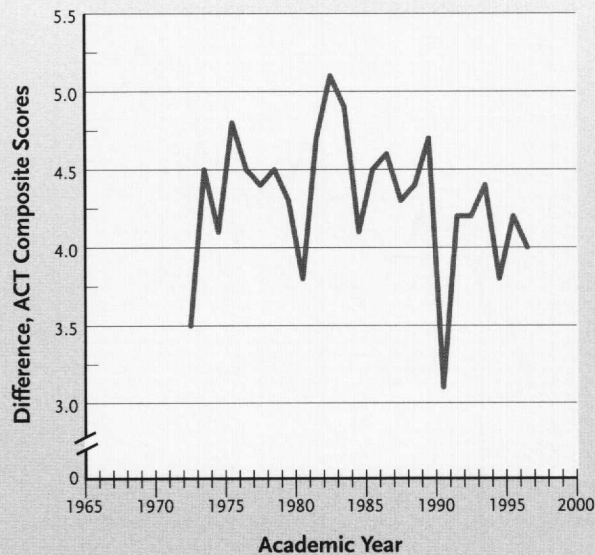


FIGURE 3: DIFFERENCE BETWEEN MEAN ACT-COMPOSITE SCORE AND MEAN OF NATIONAL COLLEGE-BOUND ACT-COMPOSITE SCORES, BY ACADEMIC YEAR

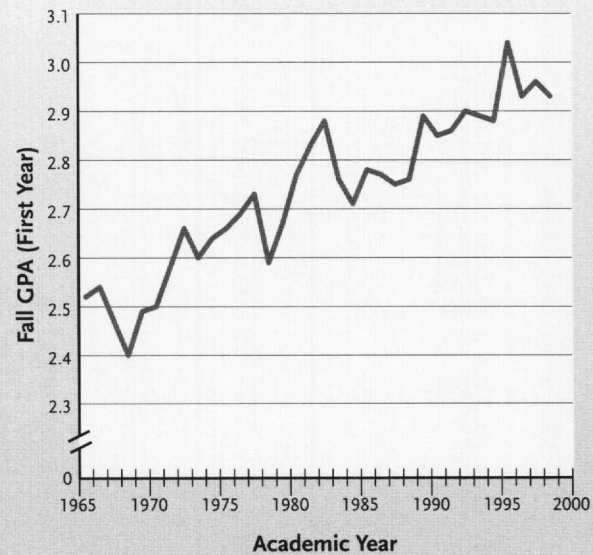


FIGURE 4: FALL TERM GPA FOR FIRST-YEAR STUDENTS, BY ACADEMIC YEAR

under study and even though different people have taught calculus from year to year, the course content and standards for student performance have remained remarkably consistent.

Basic research revealed that with the exception of one class in 1983, the proportion of each cohort enrolled in Calculus I was quite constant at between 25 and 30 percent. When the mean grades for Calculus I were plotted, it became evident that something else was quite constant, too. In spite of some fluctuation, the mean GPA for MTH 121 had hovered at or about the 2.5 threshold since 1969. Moreover, the proportion of high grades ("A," "AB," and "B") given in Calculus I (Figure 7) fluctuated at or about the 50 percent level, intersecting with the overall trend line for that ratio (see Figure 5) only five times in 30 years. Clearly, calculus had escaped the inflationary spiral! Perhaps the degree of stability in course content, level of expectations, and teaching philosophy had contributed to the lack of an upward trend in the calculus grade pattern.

Discussion

Using the Bejar and Blew model, the data presented above gave reason to believe that grade inflation, in general, had occurred. The mean ACT-Composite score, after accounting for changes in scoring, had remained essentially the same at about 24. Grades earned by students in Calculus I averaged in the neighborhood of 2.50 throughout the 30-year period. Nonetheless, all other indicators considered here—first-term GPA, the proportion of "A," "AB," and "B" grades given to the entire campus each year, and the median graduate GPA—increased steadily and significantly. Other related indicators—the proportion of students on the dean's list and the proportion of graduates earning Latin honors—increased accordingly. Not surprisingly, these

local trends followed trends at other institutions reporting data for comparable categories, but why?

In some ways, it was easier to explain why grade inflation should not have occurred. Some research has attributed higher grades to changes in the grading scale or changes in the policies related to class load and class withdrawal. The college in this study had not changed either policy since the early 1970s (about three years into the study). Even if these factors had contributed to grade inflation at that time, the impact should have plateaued by the middle of that decade. Another factor that may have been related to higher grades was change in curriculum. Both general education and major requirements became increasingly rigorous during the period under study.

In 1974, the faculty increased general education requirements for each of the three academic divisions from two courses to three. In the early 1980s, the faculty added a requirement for history and for fine arts to the existing list of nine courses. Then, in the mid-1990s, the faculty added a requirement for a mathematics course to the science requirement and required a year of foreign language for all students who had not completed a foreign language sequence in high school. Thus, both the number of general education credits required and the degree of difficulty for general education were increased.

Similar changes took place in many major fields. In the early 1980s, for example, the major in physical education (teaching and coaching) was transformed to become a kinesiology major. A number of majors added cognate requirements (*e.g.*, one year of foreign language study for English and a semester of advanced composition for foreign language). These changes raised the bar with regard to rigor in each of the programs listed and should have had the effect of containing grade inflation, not contributing to it.

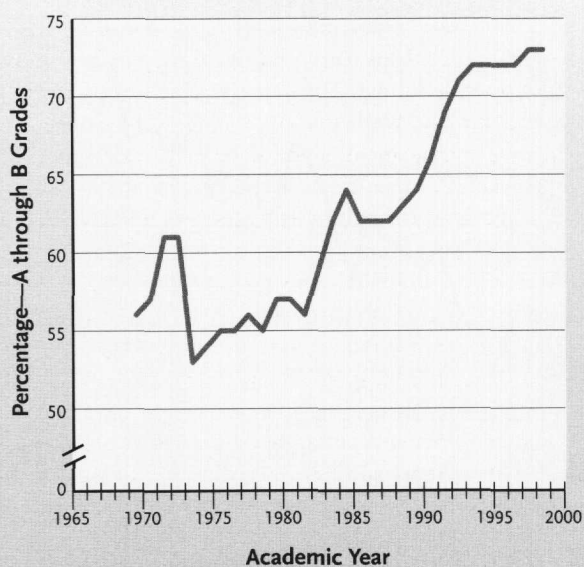


FIGURE 5: PERCENTAGE "A" THROUGH "B" GRADES, BY ACADEMIC YEAR

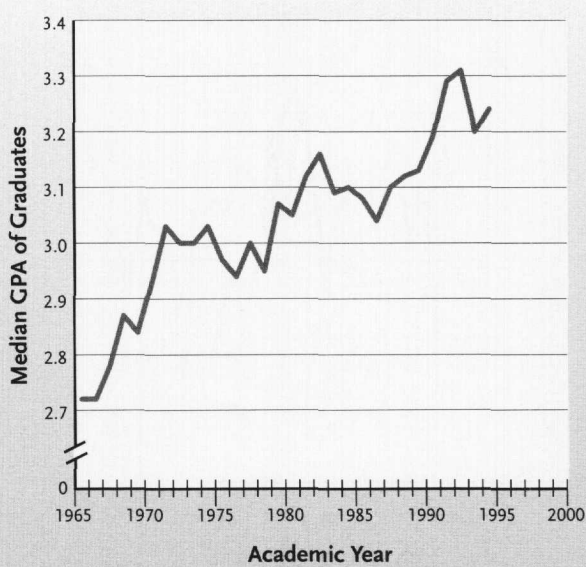


FIGURE 6: MEAN GRADE POINT AVERAGE, BY GRADUATING CLASS (ACADEMIC YEAR)

On the other hand, a number of likely suspects may have contributed to an inflationary spiral. Increased costs and increased dependence upon financial aid during the past 30 years may have affected student motivation to achieve grades high enough to remain eligible for scholarships. Moreover, rising expectations among students and parents may have contributed to higher grades. Since an increasingly higher proportion of the students who enrolled earned high school GPAs above 3.5, it is reasonable to expect that they would aim for the same level of achievement at the college level. The problem with expectations, however, is that they are nearly impossible to measure. It may also be the case that higher standards on campus have contributed to grade inflation. For example, the teacher education program raised the GPA requirement for admission to student teaching from 2.25 to 2.50 in the 1980s and from 2.50 to 2.75 in the 1990s. These changes paralleled significant increases in the proportion of students seeking teacher certification. Similarly, when the faculty changed the criteria for the awarding of Latin honors at graduation from a minimum of 3.25 to 3.40 for cum laude (3.6 for magna cum laude and 3.8 for summa cum laude), the proportion of graduates earning Latin honors declined for the first two years after the change of criteria, but soon increased and eventually surpassed the former rate. In both cases, it could be that students worked to achieve the necessary level of grade performance in order to accomplish career goals and honors that they deemed important.

Two demographic factors seemed to be possible contributors as well. First, it was noticed that the increase in grade point indicators paralleled the increase in the proportion of women students. (This group comprised 46 percent of the student body in the early 1970s and 56 percent in the mid 1990s.) This relationship was examined further and eliminated as a possible fac-

tor when a comparison of the increases in college GPA for men and for women were determined to have been about the same. A second factor involved the turnover in faculty, which occurred during the last ten years of the period under review (a period of significantly higher grades). More than half of the faculty in 1998 had been hired after 1989. In addition, the proportion of sections taught by part-time faculty increased from 10 percent to 19 percent in the same time period. The new and adjunct faculty were more diverse. Many of them were just out of graduate school and were accustomed to graduate school grading norms. They were untenured and, to acknowledge the speculation with which this paper began, may have felt vulnerable to student evaluations. And, the new full-time faculty were quite interested in their respective research agendas. In other words, they had a scholarly life away from the classroom. It is this area where the next level of research must occur.

One, more positive or uplifting, host of changes may help to explain the higher grades since 1969. The college had spent a great deal of time, money, and energy in efforts to improve teaching and provide support services to students. The interest in pedagogy is natural at an institution where good teaching is valued. The college sponsors workshops on campus to help faculty develop new pedagogical techniques, seeks grant support to help individual faculty improve their methods, and makes teaching a focal point of ongoing discussion in the faculty via a campus newsletter about teaching and a faculty forum program by which faculty share their research and pedagogical interests. Moreover, the faculty have changed the way they give essential information to students over the years. Syllabi with a standardized structure are required for all classes now. Introductory courses must, by rule, offer students two evaluations prior to mid-term instead of just mid-term and final exams. Instructors are encouraged to utilize special low grade reports to give early warning to students who are having difficulty. Mid-term grades are collected and mailed for all first- and second-year students. (Mid-term grades for sophomores were added in the late 1980s.)

Similarly, over the years, the college has invested heavily in making sure that students receive adequate support in the form of summer transition programs; tutoring (including special programs in English and mathematics); special workshops for students in time management, study skills, and test-taking; and special programs to support students with ADD and/or learning disabilities. Further, the college has provided training for new advisors since 1989. These workshops were designed to help advisors schedule students in ways which avoid the traps of missed pre-requisites, problematic schedules (*e.g.*, all lab, all lecture, all reading course), and staying enrolled in programs which are destined for failure (*e.g.*, pre-engineering students who fail calculus and/or physics).

The impact of "better support," however, is extremely difficult to assess. This is especially the case when the treatments are made without careful pre- and post-measurements, when the treatments overlap, and when the treatments are provided by different offices on the campus. However, because each of these treatments was added with the express intent of improving student classroom performance, it should be no surprise that

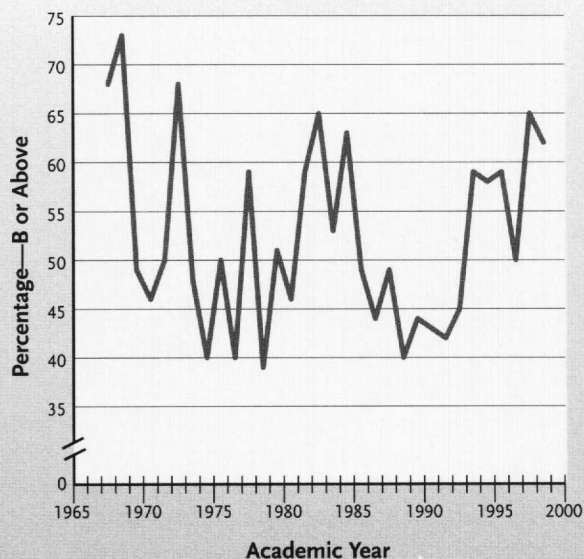


FIGURE 7: PERCENTAGE OF "B" OR ABOVE GRADES IN FIRST-TERM CALCULUS IN FALL SEMESTER, BY ACADEMIC YEAR

grades have steadily improved during the past three decades. Still, it leaves open to debate the relative absence of change in calculus grades. Perhaps some subject matter is impervious to the treatments described above.

Summary

And so the inflation enigma continues. As teachers, faculty constantly strive to find ways to help their students achieve at a higher level. Yet, once there is measurable evidence that progress may have been made, many faculty immediately become suspicious and assume that some force (e.g., student evaluations) is mysteriously causing standards to erode and quality to suffer. In this case, the quality of entering students seemed to show continuous improvement by virtue of higher high school grade point averages and slightly improved test scores. The college had raised admission standards two different times during the three decades of the study, and the proportion of honors students soared while the proportion of at-risk students dwindled. Moreover, the college initiated numerous programs to improve teaching, to support students, and to provide students with timely feedback about their learning experiences. In addition, the faculty had raised standards for general education, admission to teacher education (twice), and graduation honors, and it appears that the students have risen to these challenges. Because higher grades were expected to follow from these changes and because the changes occurred over the entire period of the study, it should have been no surprise (and certainly not an unwelcome one) that students earned higher grades (except in calculus). If there is any conclusion to be drawn from this examination of the grade inflation issue, it is this: simply attributing higher GPAs, more "A" and "B" grades, and higher proportions of honors students to inflated grades is not productive. The example provided by this case study reveals how complex the issue of grading can be and provides a framework for careful consideration of this issue on other campuses.

References

- ACT, Inc. *ACT Class Profile Reports, National Norms, 1972-1994*.
- Bejar, Isaac, I. and Edwin O. Blew. 1981. Grade inflation and the validity of the Scholastic Aptitude Test. *American Educational Research Journal*. 18(2): 143-156.
- Cohen, Henry. 1984. Inflated grades, deflated courses: Has insecurity induced compromise? *Change*. May/June: 8-10.
- Geisinger, Kurt F. 1979. A note on grading policies and grade inflation. *Improving College and University Teaching*. 27(3): 113-115.
- Goldman, Louis. 1985. The betrayal of the gatekeepers: Grade inflation. *The Journal of General Education*. 37(2): 97-121.
- Grossnickle, Donald R. and William B. Thiel. 1979. Grade inflation—who owns the guilt? *NASSP Bulletin*. 63(29): 1-5.
- Kolevzon, Michael S. 1981. Grade inflation in higher education: A comparative study. *Research in Higher Education, Journal of the Association for Institutional Research*. 15(3): 195-212.
- Lacher, David A. and Steven M. Wagner. 1987. MCAT scores to grade point ratio: An index of college grade inflation. *College & University*. 62(3): 201-206.
- Millman, Jason, Simeon P. Slovacek, Edward Kulick, and Karen J. Mitchell. 1983. Does grade inflation affect the reliability of grades? *Research in Higher Education*. 19(4): 423-429.
- Nelson, Jon P. and Kathleen A. Lynch. 1984. Grade inflation, real income, simultaneity, and teaching evaluations. *Journal of Economic Education*. 15(Winter): 21-37.
- Pedersen, Daniel. March 3, 1997. When an A is average: Duke takes on grade inflation. *Newsweek*. 129(9): 64.
- Potter, William P. 1979. Grade inflation: Unmasking the scourge of the seventies. *College & University*. 55(1): 19-26.
- Summerville, Richard M., Denise R. Ridley, and Terry L. Maris. 1990. Grade inflation: The case of urban colleges and universities. *College-Teaching*. 38(1): 33-38.
- Wegman, Jack R. 1987. An economic analysis of grade inflation using indexing. *College & University*. 62(2): 137-146.
- Weller, L. David. 1986. Attitude toward grade inflation: A random survey of American colleges of arts and sciences and colleges of education. *College & University*. 61(2): 118-127.
- Ziomek, Robert L. and Joseph C. Svec. 1995. High school grades and achievement: Evidence of grade inflation. *ACT Research Report Series*. 95(3): 16.