DOCUMENT RESUME

ED 298 869 HE 021 860

AUTHOR Summerville, Richard M.; And Others
TITLE Grade Inflation in the Eighties: The Case of Urban Colleges and Universities. AIR 1988 Annual Forum Paper.
PUB DATE May 88
NOTE 28p.; Paper presented at the Annual Forum of the Association for Institutional Research (28th, Phoenix, AZ, May 15-18, 1988).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Comparative Analysis; *Departments; Differences; *Grade Inflation; *Grade Point Average; *Grading; Higher Education; *Institutional Characteristics; Institutional Research; Intellectual Disciplines; Surveys
IDENTIFIERS *AIR Forum; Christopher Newport College Va

ABSTRACT

A study of grading practices at Christopher Newport College (Virginia) was conducted, using as data the grades awarded from 1979 to 1986. Differences between "in-department" and "out-of-department" grades were also analyzed, for all students enrolled in courses in more than one department. Results included the following: (1) the average grade of the semester just completed was 2.70, higher than the official designation of the average grade; (2) there were pronounced differences in grading among departments; (3) average grades increased during the period studied; and (4) analysis of departmental differences does not support the departments' claim that superior students in a given department account for the differences. An inter-institutional study, now ongoing, was then undertaken. A total of 351 peer institutions was surveyed with a response rate of 35%. Preliminary results from the ongoing analyses are discussed. Among the overall conclusions is that departmental and/or disciplinary differences, transcending institutional boundaries, affect grade point average elevation considerably. Data analyses are appended in tables and figures. Contains 4 references. (KM)

GRADE INFLATION IN THE EIGHTIES:

THE CASE OF URBAN COLLEGES AND UNIVERSITIES

Richard M. Summerville

Dennis R. Ridley

Terry L. Maris

Christopher Newport College

Paper presented at the Forum of the Association for Institutional Research, May 18, 1988, Phoenix, Arizona.

for Management Research, Policy Analysis, and Planning

This paper was presented at the Twenty-Eighth Annual Forum of the Association for Institutional Research held at the Hyatt Regency Phoenix Hotel, Phoenix, Arizona, May 15-18, 1988. This paper was reviewed by the AIR Forum Publications Committee and was judged to be of high quality and of interest to others concerned with the research of higher education. It has therefore been selected to be included in the ERIC Collection of Forum Papers.

Teresa Karolewski
Chair and Editor
Forum Publications Editorial
Advisory Committee

Abstract

Findings of a study of grading practices at an urban, non-residential college are presented. They reveal a general leveling of grade inflation from earlier years, with great variation among departments within the college. Other data identify departments/ disciplines as inflationary or deflationary relative to others. A survey designed to explore the generalizability of findings to similar institutions nationwide is presented together with findings and recommendations.

GRADE INFLATION IN THE EIGHTIES:
THE CASE OF URBAN COLLEGES AND UNIVERSITIES

Concern with the phenomenon called grade inflation often starts from a distressing observation that "average" grades have drifted upward, or that honors recipients are becoming more numerous than is warranted. That is basically how the phenomenon became an issue and an object of study at Christopher Newport College (CNC). However, grade inflation is more properly defined not merely as rising grades, but as "an increase in grade point average without a concomitant increase in achievement" (Bejar and Blew, 1981). This apt definition forces investigation to look for alternative explanations for higher GPA's and to test whether true grade inflation exists. One frequently heard explanation at the intra-institutional level is that, "Department X awards high grades because department X attracts exceptionally high-achieving students." This is one alternative explanation, among others, that was examined as grade inflation was studied recently at CNC. The college wanted to know whether true grade inflation existed, what was the scope of the problem if it existed, and what corrective steps, if any, ought to be taken.

The immediate problem was to deal with the perception, on the part of the Board of Visitors and others, that there were too many honors graduates. An immediate problem, particularly one perceived as such by local persons of influence, often urges upon decision-makers an immediate response. And so, in part, it happened that way at CNC. There

was an immediate adjustment of the GPA cutoff for honors graduates. Fewer people were being given honors, and so one aspect of a deeper problem lost its urgency. However, the concerned persons were also persuaded that there might be a more serious problem here, i.e. one of true grade inflation. Wisely, CNC's governing board mandated that a study be conducted to examine both the extent of such a problem at the institution and its implications.

It is important to understand this context because it has had a direct bearing on how the study was conducted, first locally, and secondly in the extension to many other institutions.

There were several concerns and questions that played a prominent role in the research, arising out of the context described above. It was recognized, for example, that at CNC, standards for consistency in the meaning of grades are stated officially and publicly, in the College catalog. While one can dispute the meaning of such terms as "excellent", "good", or "fair", a term like "average" -- which happens to be the official CNC definition of a "C" grade -- is one which can become increasingly dissonant with grading practices if true grade inflation exists, since in that case average students are awarded grades more and more above the official average. Another concern stems from the recognition that the GPA at CNC is the common coin used for such purposes as placing students on the dean's list or issuing academic warnings and placing students on probation. But what if there are significantly different grading standards among departments? What are the implications of using a number like the GPA which makes no allowances for any such differences in grading practices? These were among the concerns that motivated the local study about to be described.

The questions to be addressed by the study may be stated

succinctly:

1. Does the average grade conform to the official definitions of grades?
2. Are there significant differences among departments with regard to grading practices; if so, what are the implications for using the GPA?
3. To what extent have the grading practices changed over time? What are the trends? How strong are they? What are the implications?
4. To what extent are relatively high grade averages in a given department attributable to the ability of that department to attract "high achieving" students?
5. To what extent are higher grades in one department due to other factors unrelated to grading practices per se, e.g., the concentration of courses of lower level as compared to other departments?
6. Finally, the question of external validit and generalizability of the study suggested itself from the start. That is, to what extent are circumstances at CNC different from those at other colleges and universities across the country? This question, however, had to be held in abeyance, pending the study of the local situation.

It was discovered only after beginning the study that these same concerns had been raised years earlier at CNC, but never thoroughly or systematically studied. A severe limitation was that there were no computerized records before 1979. However, documents from the State Council of Higher Education for Virginia (SCHEV) revealed that

tabulations of the frequencies of each grade had been conducted for each public senior institution in the state; for CNC, these data extended from 1967 to 1976. As shown in Figure 1, these frequencies showed a dramatic decline in the proportion of "C" grades awarded over time, although the college persisted in calling it the "average" grade. There was a concomitant rise in the A and B, and a rather pronounced decline in D's and F's. All of this was well known in the 1960's and 1970's, of course. Translated into GPA estimates, the CNC average grade climbed from 1.81 in 1967 to 2.64 in 1976. How did the trend fare beyond 1976?

Figure 1 About Here

The scope of the study was limited to the years 1979 to 1986, with grades as the unit of analysis, but with a novel transformation of this basic unit playing an important role in the conclusions eventially reached. This transformation was the difference between "in-department" grades and "out-of-department" grades, for all students enrolled in a given department's courses who were at the same time taking courses in other departments. For example, if Department X awarded grades averaging 3.0 to students who simultaneously took courses in other departments and averaged 2.0 in those other courses, the difference would be 1.0. The reason for this transformation of data will be discussed below.

Findings and Conclusions of the Local Study

1. At the time the study was done, the average grade of the semester just completed was 2.70, a low B. Again, this contrast the official designation of the average grade as a C.

As the results of the study were discussed, it was sometimes suggested that "average" might plausibly have its reference outside of the institution. Thus a C student might be "average" in relation to some body of age-mate scholars elsewhere, but still do worse than the statistically average student at CNC. This suggestion raised the question which only a broader study might answer: Do other institutions illustrate the same disparity between their official meanings of grades and their grading practices?

2. There were pronounced differences among departments with regard to the average grades given. For example, in fall, 1986, the average grade awarded in Education was 3.35; the GPA in Mathematics was 2.23. There were many other such differences, which were quite consistent and stable over the period of the study.

However, as soon as these differences were noticed or discussed, plausible explanations could be brought forward. For example, in the comparison between Education and Mathematics just mentioned, one obvious difference between the two is that whereas Education offers only upper division courses, Mathematics offers relatively few upper division courses but has heavy traffic in lower division students. If better students tend to survive to do upper level work, this factor might explain some differences, and must be investigated. Another possibility was vigorously raised by the Education faculty as soon as these results were known, namely, that their students are simply superior students. Both of these explanations for departmental differences will be examined below. Little could be concluded here except that stable differences exist; and, again, the need for perspective from other institutions was strongly felt.

3. The college-wide trend in average grades during the period

studied was strong and upward, albeit in small increments. Figure 2 illustrates this trend. For fall semesters, the strength of the trend was indicated by a correlation coefficient of +.89, and a trend line that increases at the rate of about one-fourth of a grade point per decade. Combined with the earlier data, it was clear that the GPA at CNC experienced its strongest climb in the preceding decade and was now apparently leveling out.

4. Returning to the data transformation mentioned earlier, and the issue raised by members of the Education faculty, the reason for creating this difference score between "in-department" and "out-of-department" grades should now become evident. When the Education faculty claimed that its students receive higher grades because they are superior students, no documented evidence existed that could either refute or support that hypothesis. This spurred the idea of going to the computerized records to identify students enrolled in each department's courses who were at the same time enrolled in courses outside the department, so that "in-department" and "out-of-department" GPA's of the same students at the same time could be compared. This task (which occupied an HP-3000 for 14 hours!) was conducted for the four academic years beginning with 1982-83 and ending with 1985-86. Figure 2 provides a composite of these data for three of the four years for each department. Over each department's entry on the latter figure are three narrow bars, the left-most of which represents the 1983-84 academic year and the right-most of which represents the 1985-86 academic year.

Figure 2 About Here

This figure lends credence to the concern expressed at the outset to the effect that the GPA fails to make allowances for serious differences in grading practices among departments. It also suggests that the hypothesis, "Department X awards high grades because department X attracts exceptionally high-achieving students," is inadequate to explain such variations. For the four years studied in this analysis (only three are shown here), five departments awarded grades that were, on the average, more than 0.10 grade points higher than those earned by the same students at the same time in other departments. These departments were Accounting (ACT), Education (EDU), Physical Education (PED), Military Science (MIL), and Political Science (POS). Similarly, four departments awarded grades that averaged more than 0.10 grade points lower than those earned by the same students at the same time in other departments. These departments were Biology (BIO), English (ENG), Mathematics (MAT), and History (HIS). Now one may ask, per the hypothesis: Were the first five departments' students notable "high achievers" and were those of the latter four otherwise? The following table has been constructed to suggest the answer to this question.

Table 1 About Here

Taking one department for an illustration, the Accounting Department, in 1985-86, awarded 1998 grades to students who, at the time, were also earning grades in other departments at the college. The combined GPA for within-department courses was 2.94; the combined GPA for outside-department courses was 2.78. Thus, students enrolled in this department during 1985-86 earned grades within the department that

were slightly higher than the same students earned in other departments at the same time. Similarly, a number of departments were consistently either inflationary or deflationary relative to other departments. Observe how these results weaken the argument that higher grades reflect superior achievement. Even in one case (Education) in which outside-department grades were somewhat higher than the college mean (2.86 vs. 2.70), the superiority observed was not commensurate with the elevation of those students' within-department grades relative to the rest of the college. The obverse of this phenomenon is found in the latter four departments. These data suggest that true grade inflation (and grade deflation) probably exists, or at least cannot be ruled out by the alternative hypothesis that high-grading departments acquire that status simply because they attract high-achieving students.

5. The final issue addressed in the local study was whether the relative concentration of lower or upper division enrollments could account for some of the differences among departments in overall GPA's. A significant effect of what might be called the "curricular composition" of a department ought to result in a significant correlation between the ratio of upper division to total enrollments and the differences between in-department grades and out-of-department grades for a given department. Such correlations were computed for the academic years 1982-83 through 1985-86. These correlations ranged between .26 and .33 and averaged .29. This indicates that, on the average, only about 8.5 percent of the variance in this index of grade inflation-deflation can be explained by the relative mix of upper and lower division enrollments. It seems likely, therefore, that at CNC this factor is probably not a significant one or worthy of much consideration.

Overview of Local Study as a Preparation for Broader Study

The study of grading practices at CNC led to some rather strong conclusions, but at the same time left in suspension the question of generalizability. From a practical and theoretical standpoint, further study was needed to either establish or reveal the limitations of local conclusions. Another motivation stemmed from the local response to the study at CNC. The issues being examined turned out to be ones which elicited strong emotional, and often negative, reactions. Despite assurances to the contrary, many in the academic community reacted as though the quality of their academic programs were the issue. However, the concern that started the inquiry had no bearing on the academic quality of programs, but only with the meaning of grades as measures of student achievement and with the uses to which the institution puts these measures. Some students also questioned whether the inquiry into grade inflation, either on a college-wide or departmental level, impugned the integrity of their degrees. In all of this, the need to elevate the discussion beyond the local context seemed evident, not the least important reason being to defuse a debate that had confused (in the minds of some members of the academic community) the real issues involved.

Inter-institutional Study

Although grade inflation is a well-documented phenomenon in the 1960's and 1970's, little data exists to project the trend nationwide to the mid-1980's. Juola (1987) has communicated that there is no current information which is similar in scope to his (1979) national survey of 361 graduate degree-granting institutions; further investigation also

confirms this conclusion. Such data as are available generally are limited in that, like the study described above, they fail to relate the phenomenon to the type of institution, the nature of student populations, and the like. In view of such changes as the decreasing proportion of traditional, 18-22 year-old, full-time students, it is desirable to obtain background data about institutions in order to assess, if possible, the effects of these variables on grading practices.

The selection of a sample of institutions for this study was made with the assistance of the State Council for Higher Education in Virginia. The selection procedure exploited the circumstance that a system of "bench-mark" peer institutions has been devised for purposes of setting salary bench-marks. Using a clustering procedure available through S.A.S. (FASTCLUS), institutions have been selected which are similar to the senior institutions on a number of variables. By extending the application of the procedure beyond the original peer institutions used for the salary studies, a list of 351 institutions was generated. Of these, 200 were selected for similarity to CNC in terms of being urban, non-residential, and other indicators of similarity. Approximately 150 were selected for similarity to three other urban, mostly non-residential institutions in the state.

Survey questions were designed to provide strictly parallel information to that obtained in the local study, along with background information as to institutional-type and student populations. In the case of most questions this was straightforward; i.e., the known or estimated overall GPA for the nearest fall term, the official definitions of grades, if any, etc. In one case, however, the only approach that could be taken was by indirection since it was not

reasonable to ask responding institutions to perform an analysis that had taken an (admittedly aging) CNC computer fourteen hours to perform. This concerned the contrast between "in-department" and "out-of-department" grading. In lieu of asking for such analysis, the survey asked whether data such as were reported were available, for documentation, and for respondent's best guesses as to differences between departments in the severity of grading relative to one another.

Findings and Conclusions of Survey Study

After two mailings, an overall response rate of 35.0% was achieved. No significant differences have as yet been found between the 116 usable surveys and data gathered from bibliographic sources on a random sample (N=80) of the non-responding institutions.

The clustering procedure used in selecting the sample permitted institutions to be included which were quite dissimilar on some variables and similar on others. However, the sample could be characterized as mostly urban or suburban institutions (76.8%); in addition, 56.0% are predominantly non-residential. Respondents characterized their student bodies as "traditional" 78.4% of the time.

The following are highlights from the analysis of survey returns, which is ongoing as of this writing.

1. The mean GPA, over all respondents for fall, 1986, was 2.71. This of course is almost exactly the same as the result for CNC in 1986. The mean grade earned by undergraduates is therefore typically a low B, and it has not varied from that standard by more than 0.10 grade points at least since 1979.

The modal definition for a C grade is "average". Of those who so responded, 100% of those also prov.ding an overall GPA have a mean

undergraduate GPA of 2.50 or higher, i.e. a low B. Therefore, as at CNC, the official definition of "average" frequently does not conform to the actual or statistically average (mean) GPA.

2. Thirty-nine institutions provided some basis for department (or in some cases division, college, or school) GPA's for the years 1979 to 1986 or some part thereof. In many cases the data were contained in thick reports sent along with the surveys. It became a challenge to reduce the complexity of these data in a meaningful fashion. Despite this complexity, the 39 institutions' data could be summarized by the statement that, on average, there is more than three times as much dispersion between departments as between years. This means that departmental GPA differences are relatively stable and consistent over time, as could be readily seen in examining the data.

Furthermore, preliminary analyses suggest that there are recurring typical differences among departments in mean elevation of the GPA. For example, Education and Fine and Applied Arts, on the one hand, and Physical Sciences, Mathematics, and Social Sciences on the other hand, typically represent high- and low-grading departments, respectively.

3. Correlations were calculated between GPA's and years for all institutions which reported five or more years of data for their institution (N=39). The range of these correlation coefficients varied between +.98 and -.97, with a mean of -.01. The distribution of these correlation coefficients is shown in Figure 3.

Figure 3 About Here

It is evident from these data that if there is any relationship between year and GPA, it is as likely to be negative as positive in this

sample. At this stage no analysis has been attempted to explore why some institutions report data showing a correlation and some do not. However, the data suffice to suggest that there is no longer any general phenomenon of rising GPA's with the passage of time, as was generally true during the 1960's and 1970's. Such trends as were found, either rising or falling, tend to be in extremely small increments or decrements.

4. The survey data were inadequate to answer the question of whether relative GPA elevation of departments can be explained as due to the ability level of the students who enroll. In the absence of evidence to the contrary, based on the local CNC study, a reasonable hypothesis is that stable differences among departments cannot be explained away as due to average abilities of the students who enroll.

Overall Conclusions and Recommendations

Of considerable interest is the finding that no general pattern of GPA increase or decrease is found in these data. An affirmative answer is suggested to Juola's (1979) question, "Grade inflation in higher education -- 1979: Is it over?" In the mid-eighties at any rate, the dependence of GPA on the passage of time appears not to be a general phenomenon among the types of institutions represented in the sample. It can also be said with greater confidence than before that departmental and/or disciplinary differences, which transcend institutional boundaries, do affect GPA elevation to a considerable degree. Such a result has been previously reported (Prather, Smith, and Kodras, 1979). However, we recommend that more studies be done which would test the generality of the conclusion that such differences are not due to selective attraction of students on the basis of ability.

References

Bejar, I.I., & Blew, E.O. (1981). Grade inflation and the validity of the scholastic aptitude test. American Educational Research Journal, 18(2): 143-156.

Juola, A.E. (1979). Grade inflation in higher education -- 1979: Is it over? Unpublished manuscript. Michigan State University, Learning and Evaluation Service, East Lansing, Michigan.

Juola, A. E. (1987). Personal Communication.

Prather, J.E., Smith, G., & Kodras, J.E. (1979). A longitudinal study of grades in 144 undergraduate courses. Research in Higher Education, 10: 11-24.

TABLE 1

COMPARISONS BETWEEN DEPARTMENTS VARYING IN GPA ADVANTAGE FOR IN-DEPARTMENT GRADES

DEPARTMENT	(A) "IN-DEPTH" GPA	(B) "OUT-OF-DEPT" GPA	(A) - (B) "DEPARTMENT" GPA ADVANTAGE	"ALL-COLLEGE" GPA
Accounting	2.94	2.78	0.16	2.70
Education	3.35	2.86	0.49	2.70
Physical Education	2.87	2.60	0.27	2.70
Military Science	3.22	2.23	0.99	2.70
Political Science	2.83	2.69	0.14	2.70
Biology	2.33	2.54	-0.21	2.70
English	2.23	2.39	-0.16	2.70
Mathematics	2.23	2.47	-0.24	2.70
History	2.34	2.50	-0.16	2.70

FIGURE 1

DISTRIBUTION OF GRADES AT CNC:

1967 - 1976

DISTRIBUTION OF GRADES

1967–1976

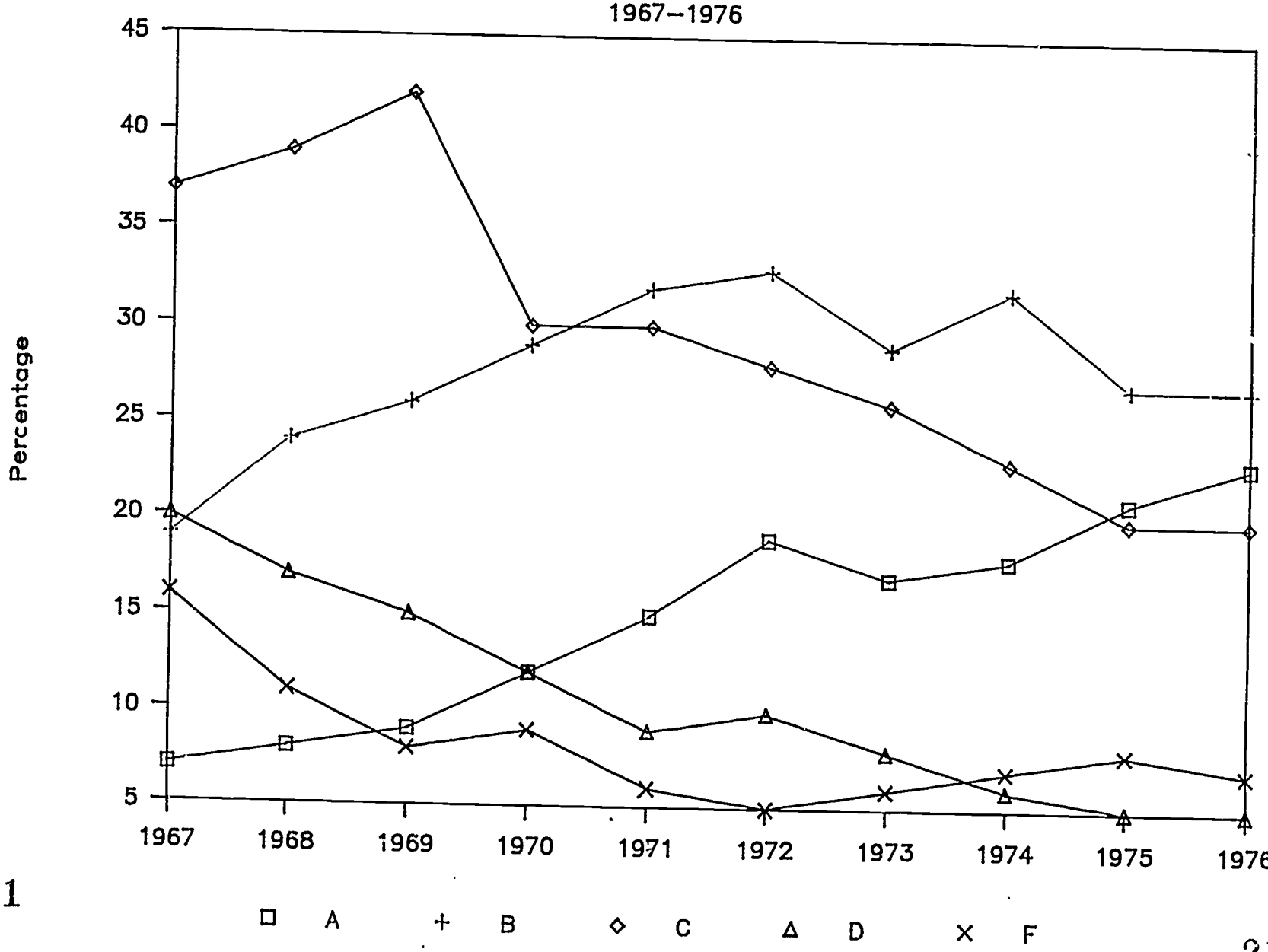

FIGURE 2

GRADE POINT DIFFERENCES BY DEPARTMENT

(IN-DEPARTMENT GPA MINUS OUT-OF-DEPARTMENT GPA):

1983-84 THROUGH 1985-86

GRADE POINT DIFFERENCES BY DEPARTMENT

ACADEMIC YEARS 83–84 THROUGH 85–86

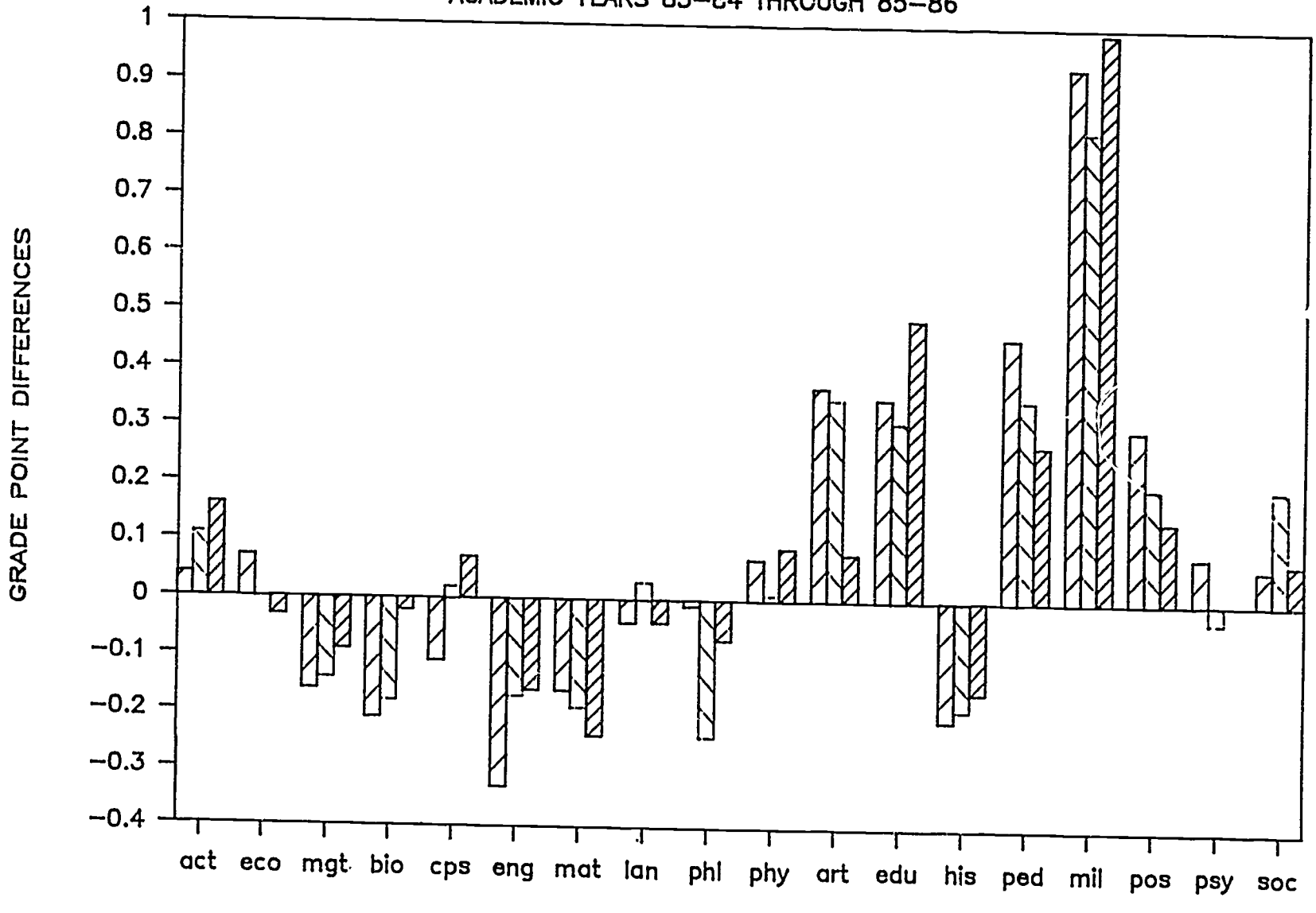

FIGURE 3

CORRELATION FREQUENCIES

(CORRELATIONS BETWEEN YEAR AND GPA):

1979 - 1986

CORRELATION FREQUENCIES

(Correlation Between Year and GPA)