Senate Committee on Examination and Grading Procedures.
Report to Senate: 31 July 1967

## Origins:

Following upon the publication and distribution of the Spring semester's examination results, the Faculty of Science expressed concern over the wide-ranging disparity of grades awarded by various faculties and departments within the University, noting specifically that the award of grades may have "important budgetary implicationsis. 1 After some preliminary discussion, Senate agreed to postpone further consideration of this question until a Committee of Senate had been able more thoroughly to investigate the issues involved, and report back to Senate under terms of a motion that "a Committee be formed to examine the examination systems and grading procedures currently in use in the University and to make recommendations to Senate. As a result of consultations between the President and his Deans, the Committee was duly established.

## Method of Approach:

The Committee held three reasonably lengthy formal sessions (the 12th, 18 th and 26 th of May) and a number of informal meetings which concluded on 28 July. It was agreed at the outset that the implications arising from the Senate directive were both far-reaching and profound, and would touch upon the essence of university teaching activity. It was further agreed, if only tacitly, that it would be exceedingly difficult for the Committee to make recommendations to Senate which could be regraded as binding upon the whole University, and at the same time, academically appropriate to the variety of disciplines taught. The Committee felt, however--perhaps only instinctively--that there did exist somewhere between the pole of anarchy and the pole of regimentation, a position bis a vis grading and examining which best reflected the independence of the individual instructor and the responsibility of that instructor to both his students and his subject. The Committee thus sought to discover whether, and in what form, such a position existed at Simon Fraser.

A questionaire was drawn up and sent to Department Heads for distribution among their faculty who had taught courses in either the Spring 1967 or Fall 1966 semester. Its purpose was to gather information about the manner in which different courses were taught and examined. Questions asked of faculty ranged from the amount of time spent in lectures, tutorials, labs etc., through the percentage of the final grade made up from examinations, essays, tutorial participation etc.,", to the role of TA's in a course. 3 The information provided by the questionaire was then tested against actual student performance. Based upon the course grade distribution figures for the spring semester and upon cumulative grade point averages (cum .GPA) of students enrolled in those courses, graphs were drawn which indicated the relationship between the level of achievement (cum. GPA), and actual performance in a specific course. Where noticeable divergencies existed between the cum. GPA of

1. For the complete statement, see Minutes of Senate, April 1967.
2. Ibid.
3. Copies of the original questionaire may be made available to members of Senate upon request.
students enrolled in a course and their subsequent final grade for that course, the appropriate questionaire could be consulted to see what kind of grading techniques had been employed by the instructor.

## Results:

Response to the questionaire was generally disappointing; 62 were returned completed: 20 from the Faculty of Science; 40 from the Faculty of Arts; 2 from the Faculty of Education. Three of the largest departments within the University--English, Modern Languages, Economics and Commerce--returned only 6 questionaires among them. With these exceptions, however, it was felt that there was still a large enough sample to go ahead with the report, although some of the conclusions must necessarily remain tentative. Selected highlights of the questionaire may be summarized as follows: (figures are on a Universitywide basis)

| 30.9\% of | the | final | grade | was | based | upon | final examinations |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 21.3\% | 18 | 1 | 11 | ${ }^{8}$ | " | 8 | essays |
| 15.2\% | " | 4 | 1 | 1 | \% | 1 | mid-terms |
| 11.0\% | " | 11 | $\because$ | 18 | 1 | 18 | tutorial participatio |

17 courses had no final examinations; 14 had no examinations at all
$54.8 \%$ ( 34 courses) reported TA association with the course
47.0\% ( 16 courses) reported TA as 'responsible for part' of final grade
$41.2 \%$ ( 14 courses) reported TA as 'advisor to professor" in determining final grade
2.9\% (1 course) reported TA as having 'sole responsibility' for final grade
$69.0 \%$ reported that grades were based upon a combination of relative and absolute standards of achievement
$\underset{6.5 \%}{24.5 \%}$ awarded grades according to an absolute standard only ${ }_{n}$
The first of the following graphs represents all courses given in the Spring semester having 20 or more students enrolled. It shows the relationship between the anticipated general level of achievement (based upon cum. GPA) and actual achievement in specific courses. The remaining graphs--one for each department, with courses from that department-illustrate essentially the same thing expressed in a different way. The solid line is the cum. GPA of students; the broken line is their actual performance in the specific course, with the point on each line indicating the mean in both cases. The lines are divided into quartiles, that is to say, $50 \%$ off the line ( $25 \%$ on each end of the two middle quartiles). Cases where the solid line and broken line (including both means) most nearly correspond are indicative of grading practices in a course which correspond to the students ${ }^{9}$ overall level of achievement as measured by their cum. GPA.
4. The complete results of the questionaire may be seen by members of Senate upon request.

To illustrate further possible implications of the relationship between grading practices and course achievement, the Committee chose four courses which identify some of the problems involved in an investigation of this kind. In view of the low response to the questionaire these courses are again not necessarily representative.

Philosophy 100 and

Physics ill:
PSA 101:
History 222:
course grade very similar to cum. GPA course grade higher than cum. GPA course grade lower than cum. GPA

Philosophy 100: There were 167 students enrolled at the beginning of the course; 152 at the end. The method of instruction was 2 hours lecturing, 1 hour of tutorial per week. There was no final examination in the course; the final course grade was based upon:
Essays: $90 \%(21.3)^{5}$
Tut. part. $10 \%$ (11.0)
The instructor reported that the final grade was the sole responsibility ${ }^{\prime}$ of the TA, yet later said that he and the TA marked the assignments. He rated the ability of his students as 'average? and described the most important aims or goals of his course as: 'knowledge of principles'; 'reasoning power'; 'critical ability'; with 'knowledge of factual detail' and 'originality, imagination' as much less important. He used a combination of relative and absolute standards (69\%) in determining his grades.

Physics 211: There were 94 students enrolled at the beginning of the course; 67 at the end. The method of instruction was 3 hours of lecturing, 1 hour of tutorial per week. The final course grade was based upon:

| Final exam: | $50 \%(30.9)$ |
| :--- | :--- |
| Tut. part. | $10 \%$ (11.0) |
| Problem sets: | $40 \%$ (no statistics) 6 |

The final exam was 3 hours and students were required to answer 6 out of 9 questions. TA's were responsible for part' ( $47 \%$ ) of the final grade. The instructor rated his students as i considerably more able ${ }^{\vee}$ than average, and described ${ }^{\circ}$ knowledge of principles, generalizations' and 'reasoning power' as the major goals of his course. He used a combination of relative and absolute standards in assigning his grades.

PSA 101: There were 310 students enrolled at the beginning of the course; 300 at the end. The method of instruction was 1 hour lecturing, 2 hours of tutorial per week. The final course grade was based upon:
Final exam: $\quad 50 \%$ (30.9)
essays: $50 \%$ (21.3)
5. Figures in parentheses represent the university-wide average.
6. The Committee felt that for some science courses 'problem sets' may be regarded as rough equivalents to essays in arts courses, hence $21.3 \%$.
gemerexy,

The final exam was 2 hours and students were required to answer 2 out of 10 essay questions. TAPs acted in an ${ }^{9}$ advisory' capacity to the instructor. ( $41.2 \%$ ) The students ${ }^{\circ}$ level of ability was not rated; the instructor regarded ${ }^{9} \mathrm{knowledge}$ of principles, generalizations' and 'originality, imagination' as the most important goals of his course, with 'knowledge of factual detail' of much less importance.

History 222: There were 90 students enrolled at the beginning of the course; 85 at the end. The method of instruction was 2 hours of lecturing, 1 hour of tutorial per week. The final course grade was based upon:

| final exam: | $50 \%(30.9)$ |
| :--- | :--- |
| essays: | $40 \%(21.3)$ |
| tut. part. | $10 \%(11.0)$ |

The final exam essay was 3 hours and students were required to answer 4 out of 11 essay questions. TAis were responsible for 'part' of the final grade ( $47 \%$ ). The instructor rated his students' ability as 'average' to ${ }^{\text {P }}$ lower than average?. The major goals of the course were ${ }^{\text {Poriginality, }}$ imagination', !critical ability', 'knowledge of principles?, with 'knowledge of factual detail ${ }^{p}$ of less importance. A combination of relative and absolute standards were used in assigning grades.

## Recommendations and Conclusions:

As suggested earlier, $\therefore$ the Committee found it difficult, if not impossible, to make specific recommendations to Senate which would be academically relevant to all--or even most--courses. A number of conclusions, perhaps very obvious ones, do emerge, however.
a) Different methods of examining and grading may all produce equally 'reliable' results, i.e. a close correlation between course grades and the cum. GPA of students. This is shown by the examples of Philosophy 100 and Physics 211.
b) It may be said that in general the 'reliability' of the course grade increases with an increase in the number and diversity of the measurements, used to establish that grade.
c) There is no obvious and conclusive relationship between the 'reliability' of a result and the role of TA's in determining the result.

More important, however, than these and other conclusions which may be drawn from the evidence of the graphs and questionaires is the general comparability of Simon Fraser grades with those of other universities. At the best of times, this is a treacherous business, but it is the kind of treachery upon which a university's reputation may ultimately hinge. Within the University the relationship between individual course grades and the cum. GPA is really not all that bad. But is the cum. GPA itself meaningful? It is, after all, made up of grades awarded in Simon Fraser courses. The Committee felt that this
was perhaps the most important question of all, and offers to Senate as one basis of comparison, some examination statistics from the University of Toronto. It should be unnecessary to remind Senate that the University of Toronto probably has the highest undergraduate admission requirements in Canada, and that the minimum as shown in the calendar is rarely adhered to in view of the large number of applicants. Clearly then, students at the University of Toronto, are academically superior to our own. The following table shows percentage distribution of grades for second year students in the University of Toronto general Arts and Science programme for the 1965-66 academic year in comparison with figures for the Spring 1967 semester at Simon Fraser.

|  | A | B | C | D | F |
| :--- | :---: | :---: | :---: | :---: | :---: | :---: |
| Toronto: | 7.7 | 31.5 | 27.7 | 16.8 | 16.2 |
| Simon Fraser: | 11.7 | 38.8 | 31.8 | 9.8 | 7.7 |

The Committee felt that this is the kind of question to which Senate ought to give very serious consideration.

W. Williams (Chairman)
I. Kendall ${ }^{7}$
E. Wells
L. Smith
31. July 1967

[^0]



$$
\sin 11 / 9 / 67 \text { Fsin, Egacuy= }
$$















Sm.119/69 Gome itwuip

PSYOHOCOGY
$$
0_{0,0}^{0} \ldots \quad \ldots \ldots
$$




[^0]:    7. The Committee would like to express its thanks to L. Kendall of the Department of Psychology for his work in preparing the statistical data in this report.
