

THE USE OF TEACHING EVALUATIONS IN MARKETING
EDUCATION: AN EXPLORATORY ANALYSIS¹

Murray L. Davis, The University of Calgary
Everett E. Johnston, The University of Calgary
Vernon J. Jones, The University of Calgary

ABSTRACT

There has been considerable controversy recently over an alleged decline in educational standards at Canadian universities. Among other issues, the widespread use of student teaching evaluations for administrative purposes has been cited as a contributing factor. This paper investigates through an exploratory analysis whether scores on teaching evaluations are positively linked to shifts in enrollment (the "shopping" effect) or to assigned grades (the "grade inflation" effect).

INTRODUCTION

The subject of this paper is the use of student evaluations of teaching performance in the administration of Business and Marketing programs and the effect upon enrollment and grading practices. The issue has become important as a result of a growing controversy over the quality of educational programs at Canadian universities. In a recently published and controversial book entitled The Great Brain Robbery, authors Bercuson, Bothwell and Granatstein, all senior professors at Canadian universities, cited declining admission standards and tenure as significant problems in declining educational quality (1984). Moreover, Bercuson et al argued that approaches to curriculum design where students can essentially shop for their subjects and widespread use of student evaluations of teaching performance by administrators have caused a reversal of responsibility for assessment "between the student who came to learn and the professor who was there to teach and assess performance" (p.81). The result is grade inflation and lowering of standards. The authors contend that the atmosphere for university education is that of a "supermarket run for the consumers' (students') benefit". (Bercuson, Bothwell and Granatstein, p.84).

Concerns raised about general university education equally apply to business and marketing programs. One of the focal points in this debate is the use of teaching evaluations and whether or not such usage is related to "shopping" for courses and grade inflation. In Canada, as in the United States, widespread use of student evaluations of teaching performances through a standard assessment form dates to the middle or late 1960's and their use has been clearly established as a tool for faculty assessment since the mid-seventies. In a study of AACSB schools reported by Seldin (1976) the percentage of schools using student ratings rose from 20.6% in 1967 to 65.3% in 1975. Student ratings were ranked third in 1975 (behind Dean and Chairman evaluations) as information utilized for evaluating teaching performance compared to ninth in 1967. Lein and Merz (1978)

later reported that over 70% of AACSB schools use student evaluation forms.

In an excellent review of the use of evaluation forms in Marketing, Wilson (1982) cited four areas of possible research: the purpose of such evaluations, instrument construction, reliability and validity. Of these, instrument construction and reliability testing seem to have been most thoroughly researched and are perhaps less controversial due to the vast amount of work in the educational psychology area. Validity issues appear to be addressed, at least in the marketing literature, with less frequency. Marks and Mollander concluded, reminiscent of Bercuson et al., that evaluation scales measure consumer "satisfaction" rather than teaching effectiveness (1977, p.368) and Houston and Davis somewhat disturbingly reported that while grade expectations were related to student evaluations there appeared to be "little relationship between knowledge gained from the course and student evaluation of the instructor" (p.195).

What is perhaps the most important issue regarding student evaluations of teacher performance is the purpose to which the results are put and the effect of this over time. Wilson (1982) states that evaluations may be used by administrators (normative) in a comparative mode, by instructors (diagnostic) to improve performance, and by students (informative) as an aid to class or instructor selection all of which imply that the instructor is likely to perceive a payoff to high ratings on student teaching evaluations. However, these presumed beneficial results of the use of teaching evaluations have gone relatively unquestioned especially for administrative purposes. Consider the recent remarks of Stoakley Swanson in the Marketing Educator:

"We also need to acknowledge that the current state of the art concerning student evaluation of faculty is still but one step removed from the reading of tea leaves and at long last to commit ourselves to doing something about it. Particularly deplorable is the present tendency to try to reduce the data from the evaluation instruments to a single summary statistic.

"This effort is made regardless of subject taught, of class level and class size, and of the number of different courses assigned to each instructor per term and throughout the academic year (not to mention any of the other multiplicity of variables at work in the dynamics of the student/teacher interface), so as to simplify and facilitate the comparing of performances across a diverse curriculum and a heterogeneous faculty. Equally absurd is interpreting student evaluation without reference to each instructor's grading practices." (Swanson, 1984)

Swanson's remarks are not a criticism of teaching evaluation as such. His remarks are not addressed

¹The authors would like to acknowledge the assistance of Gordon Barth with the data analysis.

to how the instrument is constructed nor its reliability nor even its validity if used properly. His criticism is the purpose to which the results are put and the oversimplification generally found in their interpretation. This is important because decisions are made based upon the results and these decisions are critical not only to individual careers but also the development of academic programs.

The object of this paper is to examine through an exploratory analysis the impact of teaching evaluations at a particular school where their use is a significant factor in faculty assessment. The two specific areas of concern are the relationship between teaching evaluations and enrollment or changes in enrollment (the "shopping" effect) and the relationship between teaching evaluations and grades (the "grade inflation" effect).

BACKGROUND AND DATA

This study was conducted on data compiled from teaching evaluations, course enrollment and grading at a western Canadian business school, the Faculty of Management at The University of Calgary. The Faculty was established in 1968 and grew rapidly to an enrollment of approximately 700 undergraduate students and 250 graduate students by 1978. Since then enrollment has been stable due to a quota on admissions. The Faculty uses a grade point system on a scale of 0.0 to 4.0, has a minimum admission standard of 2.3 based upon two years of undergraduate studies, and a normal range of concentrations (or majors) including marketing, accounting, management information systems, human resources management, etc. Once admitted to the Faculty, a student may choose to enroll in any major for which he or she has successfully completed pre-requisites. This study was conducted on undergraduate courses only.

Student evaluations of teaching effectiveness (herein referred to as teaching evaluations) have been used by the Faculty since its inception. However, an external review of the Faculty conducted in 1979 created a positive impetus towards improved teaching by recommending that teaching evaluations be considered more seriously in the annual review and assessment of faculty performance and that the results be more widely disseminated. Since 1981, when a new assessment instrument was introduced, teaching evaluations as administered by the Faculty with the cooperation of students and instructors have become the single most important element in the assessment of teaching performance. New course development, supervision of graduate and senior undergraduate student projects, course coordination and other factors are recognized but these are by no means central to the assessment of teaching performance. High instructor ratings as determined by students are a necessary and often sufficient condition to be assessed a good and effective teacher. These assessments are used extensively in the annual review of Faculty for salary adjustments, decisions on contract renewal, tenure and promotion.

The form used is a modified version of one developed by Hildebrand, Wilson and Dienst at the Centre for Research and Development in Higher Education at the University of California at

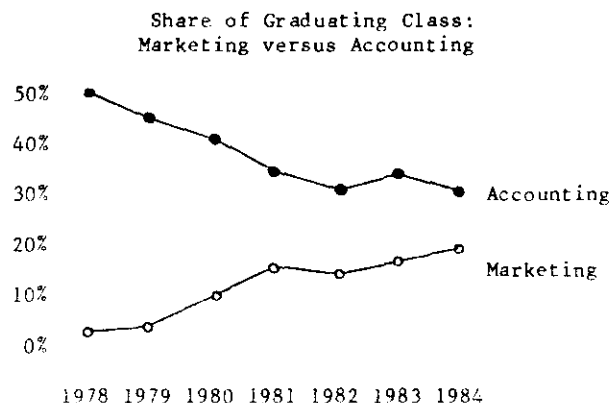
Berkeley (1971). Ratings are completed on 35 items which yield combined scores on five factors: analytic-synthetic approach, organizational-clarity, instructor-group interaction, instructor-individual interaction, dynamism-enthusiasm. Also computed are a calculated overall score which is an average of the items, an overall instructor effectiveness item, and an overall course rating item. Each item, including overall instructor effectiveness and course ratings, is scaled from 1 to 5 where 1 is described as unsatisfactory and 5 is excellent.

While data on enrollment, grades and teaching evaluations are available from 1974, the analysis was performed on data available from the Fall, 1981 to Winter, 1984 teaching terms inclusive in order to coincide with the introduction of the new instrument and change in Faculty policy regarding the use of teaching evaluations. In each term, that is each Fall session or Winter session, approximately 80 classes are offered in the undergraduate program. Over the period from 1981 to 1984, teaching evaluations were conducted in 518 undergraduate classes with enrollment greater than 5. The overall instructor and course rating scores, the average grades and course enrollment by subject area form the data base for the analysis performed in this study. The method undertaken was to examine a series of hypotheses relating scores on teaching evaluation to enrollment levels and grades.

TEACHING EVALUATION AND ENROLLMENT (THE "SHOPPING" EFFECT)

As indicated, overall enrollment in undergraduate business studies has been stable for a number of years as a result of a limitation on admissions which essentially established the program at fixed size. However, within the program itself, significant enrollment shifts can take place at the senior level by students opting to shift majors, that is, to decide to major in Marketing as opposed to something else. Prior to 1978, a Marketing program did not exist. Figure 1 plots the growth in enrollment in Marketing majors as opposed to Accounting majors since the introduction of a Marketing program in 1978.

FIGURE 1



The establishment of Marketing around 15% of the class and Accounting at 30% may simply be a stabi-

lizing in market share after the introduction of a new program. Moreover, teaching evaluations of Marketing versus Accounting courses do not show a lot of variation from each other or from the Faculty average.

To investigate the relationship between teaching evaluation and enrollment or changes in enrollment the following hypotheses were examined.

- H₁: There exists a positive correlation with course enrollments by subject with student teaching evaluations (as measured by overall instructor and course ratings).
- H₂: There exists a positive correlation between number of majors by subject with student teaching evaluations in senior courses (as measured by overall instructor and course ratings).
- H₃: There exists a positive correlation between number of majors by subject with student teaching evaluation in previous year (as measured by overall instructor and course ratings).
- H₄: There exists a positive correlation between the change in number of majors by subject with student teaching evaluation (as measured by overall instructor and course ratings).
- H₅: There exists a positive correlation between the change in number of majors by subject with student teaching evaluation in previous year (as measured by overall instructor and course ratings).

TABLE 1

Correlation of Enrollment and Student Teaching Evaluation		
Hypothesis	Instructor Evaluation	Course Evaluation
H ₁ :	.113 (.46)	.061 (.24)
H ₂ :	.181 (.74)*	.252 (1.04)
H ₃ :	-.27 (.89)	.702 ** (3.12)
H ₄ :	-.172 (-.7)	-.035 (.14)
H ₅ :	-.161 (-.52)	.589 ** (-2.31)

* () t-statistic
** significant at <.05

The only significant results are as follows (Table 1):

- 1) The number of majors by subject area is strongly correlated to the average teaching evaluation of courses in the previous year. Note that this is not the case for instructors. That is, students seem to be affected in their choice of major by their evaluation of the

course (not instructor) in the previous year.

- 2) The change in number of majors is correlated with the average teaching evaluation of the courses in the previous year. Again, note that this is not the case for the instructor rating.

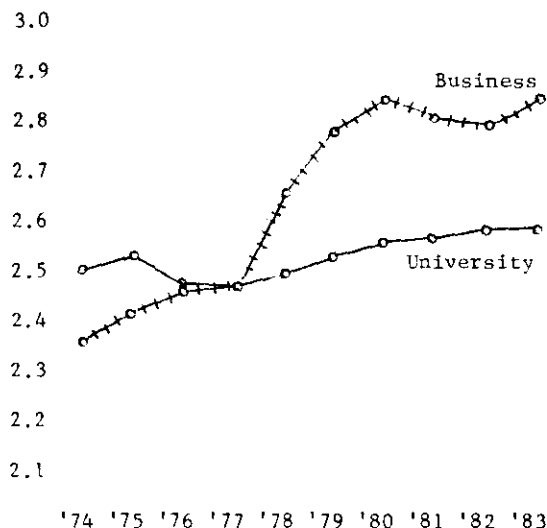
There appears no significant relationship between evaluation of instructors either in the current year or previous year.

TEACHING EVALUATIONS AND GRADES
(THE "GRADE INFLATION" EFFECT)

Figure 2 plots the average grade point at The University of Calgary from 1974 to 1983 for all undergraduate programs and the undergraduate Business program. Clearly, there is an upward trend in grade point within the Business program that exceeds the University average. However, it is difficult to conclude that this is evidence of grade inflation or the result of teaching evaluations. With the sharp increase in demand for business education throughout the seventies, there is every reason to believe that the general quality of student has increased as programs have been limited to fixed capacities. Moreover, in the past ten years, Faculty of Management undergraduate admission standards have been raised twice. It is then not surprising that the average grade point in the Faculty exceeds the University average.

FIGURE 2

Grade Point Average for Undergraduate Programs (1974 - 1983)



Nevertheless there may be areas of concern where grade inflation should be. For example, one should not expect that the average grade point for Marketing courses exceeds the Faculty average as there are no special requirements for admission to Marketing program. In this regard Table 2 is interesting as it shows differences in average grade point across subject areas over the part three years for senior level courses all of which are elective. Marketing courses show a sharp rise in

grade point average from 1981/82 and are second highest after Information Systems. Most striking is the contrast between Accounting which has experienced the sharpest decline in student numbers (see Figure 1) and other subject areas including

TABLE 2

Average Grade Point Averages - Senior Undergraduate Courses				
	1981/82	1982/83	1983/84	Average
Accounting	2.23	2.51	2.64	2.47
Finance	2.94	3.13	3.11	3.07
Information Systems	3.15	3.42	3.42	3.33
Management Science	2.93	3.08	3.15	3.04
Marketing	2.94	3.26	3.28	3.17
Human Resources Mgmt.	3.15	3.09	3.07	3.11

Marketing. Might we conclude that Marketing is softer on grades and has generated higher student enrollments and higher student evaluation through its grading practices?

In order to investigate more usefully the relationship between grades and teaching evaluation the following hypothesis was tested:

H₁: There exists a positive correlation between grade point averages across subject areas and teaching evaluation (as measured by overall instructor and course ratings).

This relationship was tested correlating average teaching evaluations with grade point averages for 1981 to 1984. The results are presented in Table 3:

TABLE 3

Correlation of Average Grade Point and Teaching Evaluation		
	Instructor Score	Course Score
Accounting	.264	.185
Finance	.473 *	.452 *
Information Systems	.173	.195
Management Science	.438 *	.510 *
Marketing	.401 *	.408 *
Human Resources Management	.108	.234

* Significant at <.05 level

It is apparent that there is a positive relationship between grades achieved in Finance, Management Sciences and Marketing with teaching evaluations for both instructor and course. This relationship is much weaker or is not significant in Accountancy, Human Resources Management and Information Systems.

CONCLUSION

The discovery of positive relationships between variables should allow for an underlying construct, for example, the possibility that higher

teaching evaluation and higher grades might both be explained by more effective teaching. Moreover, the fact that higher course evaluations in the penultimate year are related to change in choice of major is not surprising and not necessarily indicative of a decline in educational standards. Students may be quite logically choosing to major in areas where they rate course content most highly. The failure to find any relationship between enrollment and evaluation of instructors suggests students are well able to distinguish course content and the effectiveness of the teacher.

Nevertheless, if one looks carefully at the Marketing area we can observe not only a significant relationship between course evaluation and choice of major but also a positive correlation among instructor, course evaluation and grade point average. The potential organizational impact of this should give cause for concern. The important point is to recognize the limitations of such measures and to merge their use with judgement about true effectiveness in teaching and maintenance of educational standards.

REFERENCES

- Bercuson, D.P., R. Bothwell, and J.L. Granatstein (1984), The Great Brain Robbery: Canada's Universities on the Road to Ruin, Toronto: McClelland and Stewart.
- Hildebrand, M., R.C. Wilson, and E.R. Dienst (1971), Evaluating University Teaching, Centre for Research and Development in Higher Education, University of California (Berkeley).
- Houston, F.S. and J.L. Davis (1977), "Learning and Grade Expectations: A Study of Instructor Evaluations," Journal of the Academy of Marketing Science, 5, (Summer), 195-202.
- Lein, C.D. and C.M. Merz (1978), "Faculty Evaluation in Schools of Business", Collegiate News and Views, 21.
- Marks, R.B. and J.D. Mollander (1977), "Student Evaluations - A Measure of Validity in the Introductory Marketing Course," Proceedings: America Marketing Association Educators Conference, 366-69.
- Seldin, Peter (1976), "Current Practices in Evaluating Business Faculty", AACSB Bulletin, 12, (April), 1-6.
- Swanson, Stoakley W. (1984), "Academic Workstyles Reaction - time to face reality in teaching, research and service", Marketing Educator, 3, (Spring), 7.
- University of Calgary Fact Book 1983-84, (1984), 174-77.
- Wilson, Terry C. (1982), "Student-Faculty Evaluation Forms in Marketing: A Review", Journal of Marketing Education, (Spring), 7-14.