Relationship Between Student Characteristics And Their Ratings of Professors

C.E. Stufflebeam

Abstract

Student ratings were collected from 896 students over an eight year period in an introductory course in animal agriculture. The possible effects of year, season, semester, class level, major, expected grade, GPA and sex of the student were examined. Generally, student ratings were affected significantly only by sex of the student, expected grade and GPA. Women students within each expected grade level and overall rated the professor more highly than did men students. With respect to the effects of grades, the higher the expected grade, the higher the student ratings. Students with extremely high GPA rated the teacher higher; those with the lowest GPA rated him lowest. However, GPA did not seem to be a factor within the broad range of 2.0 to 3.5.

Introduction

Student ratings of instructors are in widespread use in higher education throughout the United States and Canada. Information from these ratings is being used in personnel decisions, for self improvement by professors, and by students in course selection. A considerable volume of evidence supporting the validity and reliability of student ratings as part of a system of evaluation of instructors and instruction has been cited in reviews by Aleamoni and Hexner (1980), Braskamp **et al.** (1983 and 1984), Feldman (1977) and McKeachie (1979). Many of these findings were summarized in a recent article (Stufflebeam, 1987) prepared in conjunction with the present study. Conclusions of that review were generally that the literature was quite supportive of student ratings as reflectors of student attitudes. As a group, students seem to be very perceptive and reliable judges, and are discriminating and consistent in their judgments of instructors and instruction.

Methods

This study was conducted over an eight year period during 14 semesters, and involved 896 students. The first phase of the study involved 395 students during six semesters from the fall of 1976 through the spring of 1979. The second phase began in the fall of 1979, ended in the spring of 1984, and included 501 students.

During the last week before the final examination in an introductory course in animal science, students were asked to anonymously rate the teacher. In the first phase of the study, the rating instrument contained 15 items each of which could be rated with A, B, C, D or F. As suggested by Masters (1974), five levels of response were chosen to allow students ample opportunity to discriminate, and to increase reliability of the instrument. Responses to the 15 items on the rating form were assigned values of zero through 4 with the "A" response receiving the 4. The values were summed to obtain an overall rating from each student. The highest possible rating was 60.

Items on the rating instrument asked for student opinions on such things as ability to stimulate interest, ability to communicate at the students' level, delivery of material in an interesting way, use of humor, fairness in testing, grading as well as dealing with students, knowledge of the subject, encouragement of class discussion, avoidance of annoying mannerisms, study guides plus class exercises, and overall rating compared to other teachers. Ideas for the design of the instrument were obtained from a number of rating instruments including those used by Cushman and Tom (1976), and Aleamoni (1976). In addition, information was obtained about the students' major, class level, GPA and expected grade. Students had previously been provided with information that would permit them to make valid judgments as to their grade expectations.

For the second phase of the study, the instrument was revised, increasing the number of response items to 20. Some of the questions added related to the ability of the teacher to stimulate students at several levels of learning including factual knowledge, principles and concepts, problem-solving and decision making, and creative and imaginative skills. Ideas for revising the instrument were obtained after reviewing the rating instrument used by Pandya and Curtis (1979). On the phase two instrument, a response item indicating the sex of the student was added. Before this time, the number of women in the course was too low to justify requesting this information. In the second phase, values of 1 through 5 were assigned to each of the 20 responses, with an overall maximum possible rating of 100.

Phase I data was classified into groups by semester, class level, major, expected grade and grade point average. Phase II data was categorized by year, season (fall versus spring semesters), class level, major, expected grade and sex. Comparisons among the several categories within each phase of the study were made by analyses of variance (Snedecor and Cochran, 1980).

Results and Discussion

Mean scores for the six semesters in Phase I showed differences significant at the five percent level of probability. However, the magnitude of variation seemed to be due to one unexplained low mean that

Stufflebeam is a professor of Animal Science, Southwest Missouri State University, Springfield, MO 65804-0094.

occurred during the fourth semester of the study (Table 1). Small differences among the eight semesters of the second phase were not statistically significant. Overall means and standard deviations for the two phases of the study were 50.6 ± 7.1 and 85.3 ± 8.3, respectively.

Table I. Student Rating Scores by Semester, Class, Major, Grade Point Average and Expected Grade (Phase I, 1976-1979).

Semester*	Score	No.
F 76	50.4	109
S 77	51.4	68
F 77	50.1	52
S 78	48.1	50
F 78	51.1	66
S 79	52.2	50
Class	**Score**	**No.**
Freshmen	50.2	167
Sophomore	51.2	106
Upperclass	50.7	122
Major	**Score**	**No.**
Animal Science	51.7	79
Agricultural Bus.	49.5	76
Other Agriculture	51.2	109
All Other Majors	50.1	131
Grade Point Ave.**	**Score**	**No.**
3.5-4.0	53.6	31
3.0-3.4	51.2	71
2.5-2.9	50.3	136
2.0-2.4	50.8	126
Below 2.0	47.1	31
Expected Grade***	**Score**	**No.**
A	54.2	53
B	50.9	154
C	49.9	137
D/F	47.8	51
Mean and Total	50.6	395

* Differences significant at 0.05
** Differences significant at 0.005
*** Differences significant at 0.001

Three class levels (freshmen, sophomore and upperclass) were compared in both phases of the study for their possible effects on instructor ratings. No significant differences were observed in either phase of the study (Tables I and II). Results of other studies were inconsistent in revealing a relationship between class level and student ratings. In one study of over 4000 students in 87 classes, Rayder (1968) concluded that ratings were not substantially related to class levels of the students. In a review by Feldman (1977), 16 references reported no relationship, 13 reported positive relationships, and five reported negative relationships. A review by Aleamoni and Hexner (1980) cited eight investigators who reported no significant relationships, while 18 others reported that graduate and upper division students tended to rate teachers higher than did students in lower division courses.

Table II. Student Rating Scores by Major and Class, (Phase II, 1979-1984)

Major	Score	No.	Class	Score	No.
Animal Science	86.4	148	Freshmen	85.0	275
Agricultural Bus.	84.6	136	Sophomore	85.3	135
Other Agriculture	85.1	108	Upperclass	86.3	91
All Other Majors	84.9	109	Mean & Total	85.3	501

In both phases of the study, ratings of students were classified into four majors and compared: animal science (including pre-veterinary students), agricultural business, other agriculture majors, and all other majors. In both phases, animal science majors tended to rate the professor highest and agricultural business students rated him lowest; however, the differences were not significant at the five percent level of probability (Tables I and II). These results are consistent with reports from the literature, Rayder (1968), and Null and Nicholson (1972). Some references indicated that students with a greater interest in a course tended to rate the instructor higher, Feldman (1976), and Wessel and Grewal (1982). Assuming that students within a certain major might be more interested in that area than non-majors, and if greater interest in a subject results in higher ratings, students within the major of the course might be expected to give higher ratings to the instructor. Although it was not statistically significant, this tendency occurred during both phases of the study.

Student ratings were separated into four groups according to the grades expected in the course. Very few students indicated that they expected an F in the course, so their ratings were included with the D group for evaluation. In both phases of the study, students expecting a higher grade in the course tended to rate the instructor higher. Mean scores for each grade category are shown in Tables I and III, and graphically in Figure I. The differences were significant at the 0.001 and 0.005 levels of probability for the first and second phases of the study, respectively.

These results are consistent with a majority of the reports in the literature that suggested a positive correlation between expected grades and ratings (Bradenburg *et al.*, 1977; Elmore and Pohlmann, 1978; Lockwood *et al.*, 1987; Wessel and Grewal, 1982; and Weaver, 1960). Aleamonie and Hexner (1980) cited 22 studies reporting zero relationships and another 28 reporting significant positive relationships. After reviewing over 100 references, Feldman (1976) concluded that anticipated or actual grades are positively related to instructor ratings. An earlier report by Anikeef (1953) indicated that students in lower division courses were more likely to let their expected grades influence their instructor ratings.

McKeachie (1979) suggested that the relationship between higher grades and higher ratings is not

Table III. Student Rating Scores by Sex and Expected Grade. (Phase II, 1979-1984)

Expected	Men		Women		Means*	
Grade	Score	No.	Score	No.	Score	No.
A	85.6	53	89.3	37	87.1	90
B	85.1	159	87.4	76	85.9	235
C	83.6	107	84.5	50	83.9	157
D/F	80.9	15	83.0	4	81.3	19
Means*	84.5	334	86.9	167	85.3	501

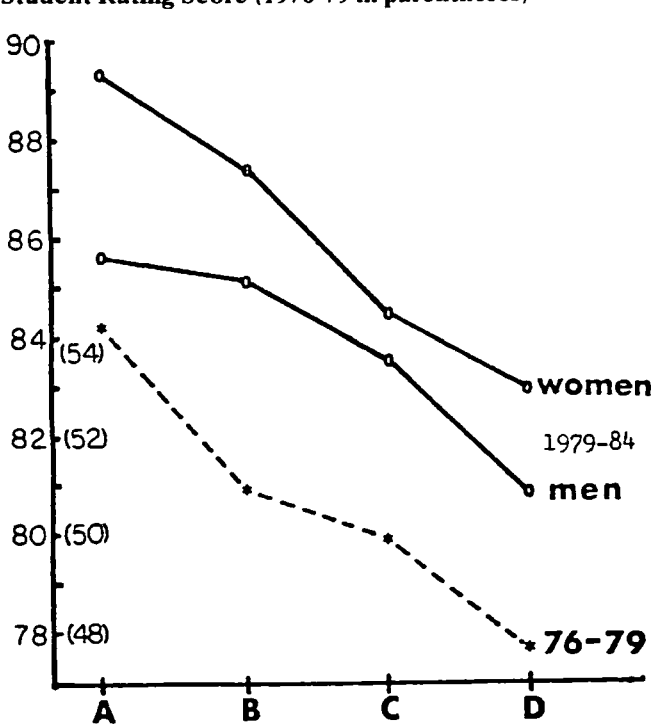

Figure 1. Effect of Sex and Expected Grade on Student Ratings of an Instructor.

necessarily a sign of invalidity of student ratings. It could be argued that in courses where students learn more their grades should be higher, and therefore their ratings should be higher. He concluded that within the normal range of variation usually observed, there probably need not be much concern about grading standards and student ratings. In an experiment controlled for student achievement, Palmer *et al.* (1978), found no effect of severity of grading on student ratings.

There was a general tendency, as expected, for students who expected higher grades to also have the higher grade point averages. Since students expecting higher grades rated the teacher higher than students receiving lower grades, it would seem logical to expect students with higher ovall GPAs to also rate him higher than students with lower GPAs. Generally this was the case, but the effect was not as dramatic as that of expected grade. Students above 3.5 rated the professor higher, while those with GPAs below 2.0 rated him lowest. Differences were significant at the 0.005 level of probability. However, there did not seem to be any differences in ratings by students within the broad GPA range of 2.0 to 3.5. This latter group represented over 84 percent of the sample. Other reports from the literature seem to indicate little or no relationship between GPA and rating scores (Rayder, 1968 and Feldman, 1976). Overall, in the present study, GPA seemed to be less of an indicator of how students rate their instructor than did expected grade.

Sex of the student was a variable in Phase II only. As shown in Table III and Figure 1, women students tended to rate the isntructor higher than did men students. This was true within each expected grade category. Rayder (1968) showed that sex of students had no effect on their ratings of teachers. This conclusion was supported in a review by Costin *et al.* (1971) of seven references that reported no differences in overall ratings made by male and female students. An extensive review by Feldman (1977) cited 26 references showing no relationships, and 26 showing that women students tended to rate teachers slightly higher than did men students on some rating items.

Conclusions

Results of studies of how student ratings were affected by several student characteristics, including major, course level, expected grade, GPA and sex of student, were mixed. Some factors seemed to have more impact than others. However, ratings appeared not to be seriously affected by most factors since the proportion of the total variation accounted for by them was usually relatively small.

In the present study, effects of year, season, semester, major, class level, sex of the student, expected grade and GPA were examined. Only sex of the student, expected grade and GPA seemed to have any significant affects. However, knowing that these or other factors may influence student ratings does not necessarily invalidate their use or make them unreliable. That very knowledge can provide a basis for adjustments. For example, if one class has a higher percentage of women students, and it is known that professors are rated higher by women than by men, the ratings can then be weighted to compensate for the difference.

Student ratings of teachers seem to be useful for providing feedback for self improvement and as part of the input for evaluation in personnel decisions. In the latter case, it seems appropriate to obtain ratings from several classes during more than one semester during the evaluation period. These ratings should then be evaluated and interpreted along with other sources of information by someone familiar with the various factors that might influence ratings, i.e., the department head or evaluation committee.

Literature Cited

Aleamonie, L.M. 1976. "Typical Faculty Concerns about Student Evaluations of Instruction," *NACTA Journal*, 20(1):16.

Aleamonie, L.M. and P.Z. Hexner. 1980. "A Review of the Research on Student Evaluations and a Report on the Effect of Different Sets of Instruction on Student and Instructor Evaluations." *Instructor Science*, 9:67.

Anikeef, A.M. 1953. "Factors Affecting Student Evaluations of College Faculty Members." *Journal of Applied Psychology*, 37:458.

Brandenburg, D.C., J.A. Slinde and E.E. Patista. 1977. "Student Ratings of Instruction: Validity and Normative Interpretations." *Research in Higher Education*, 7:67.

Braskamp, L.A., D.C. Brandenburg, E. Kohen, J.C. Ory and P.W. Mayberry. 1983. "Guidebook for Evaluating Teaching: Part I: Rationale and Principles." *NACTA Journal*, 27(4):29.

Braskamp, L.A., D.C. Brandenburg, E. Kohen, J.C. Ory and P.W. Mayberry. 1984. "Guidebook for Evaluating Teaching, Parts II-IV: Collecting Evaluative Information about Teaching." *NACTA Journal*, 28(1):19, 28(2):27 and 28(4):27.

Costin, F.W., W.T. Greenough and R.J. Menges. 1971. "Student Ratings of College Teaching: Reliability, Validity and Usefulness." *Review of Educational Research*, 41:511.

Cushman, H.R. and F.K.T. Tom. 1976. "The Cornell Diagnostic Observation and Reporting Systems for Student Description of College Teaching." *NACTA Journal*, 20(1):10.

Elmore, P.B. and J.T. Pohlmann. 1978. "Effect of Teacher, Student and Class Characteristics on the Evaluation of College Instructors." *Journal of Educational Psychology*, 70:187.

Feldman, K.A. 1976. "Grades and College Students Evaluation of Their Courses and Teachers." *Research in Higher Education*, 4:69.

Feldman, K.A. 1977. "Consistency and Variability among College Students in Rating Their Teachers and Courses: A Review and Analysis." *Research in Higher Education*, 6:223.

Lockwood, J.A., G.E. Moore and R.N. Story. 1987. "Relationship between Student and Professor Characteristics and Results of Evaluations." *NACTA Journal*, 31(2):33.

Masters, J.R. 1974. "The Relationship Between Numbers of Response Categories and Reliability of Likert-type Questionnaires." *Journal of Educational Measurements*, 11:49.

McKeachie, W.J. 1979. "Student Ratings of Faculty: A Reprise." *Academe*, 65:384.

Null, E.J. and E.W. Nicholson. 1972. "Personal Variables of Students and Their Perception of University Instructors." *College Student Journal*, 6:6.

Palmer, J., G. Carliner and T. Romer. 1978. "Leniency, Learning and Evaluations," *Journal of Educational Psychology*, 70:855.

Pandya, H.S. and S.M. Curtis. 1979. "College Students' Evaluations of Instructors and Courses." *NACTA Journal*, 23(4):11.

Rayder, N.F. 1968. "College Student Ratings of Instructors." *Journal of Experimental Education*, 37:76.

Snedecor, G.W. and W.G. Cochran. 1980. *Statistical Methods*, 7th ed. Iowa State University Press, Ames.

Stufflebeam, C.E. 1987. "Validity and Reliability of Student Ratings of Professors." *NACTA Journal*, 31(4)

Weaver, C.H. 1960. "Instructor Ratings by College Students." *Journal of Educational Psychology*, 51:21.

Wessel, K.L. and H. Grewal. 1982. *Longitudinal Study of Student Evaluation of a Course*. Department of Agricultural Economics and Rural Sociology, The Ohio State University.

A Profile of Women Scientists In Colleges of Agriculture

Barbara E. Cooper
and Janet L. Henderson

The representation of women in the scientific and engineering professions has been low, and the agricultural sciences have been particularly difficult areas for women to enter. In 1984, Vetter found high unemployment rates for women in the agricultural sciences at all degree levels. Specifically, in colleges of agriculture, the representation of women among agricultural scientists is low. Currently, women agricultural scientists at the 70 U.S. land grant universities comprise 4.6% of the total agricultural faculty (Henderson & Cooper, 1987).

A recent study (Henderson & Cooper, 1987) provided information on the numbers of women faculty in U.S. land grant colleges of agriculture. Seventy percent of the colleges have between one and 20 women agricultural scientists. The highest number of women agricultural scientists employed in any one college is 27. In relation to the total population of women agricultural scientists, the crop and soil science discipline has the highest number of women scientists (19%). The agricultural engineering discipline has the fewest number of women scientists (1%). There is, additionally, a low representation of women agricultural scientists in all geographic regions of the U.S.

Cooper and Henderson are assistant professors in the Department of Agricultural Education, The Ohio State University, 204 Agricultural Administration Building, 2120 Fyffe Road, Columbus, Ohio 43210, 614-292-6671.

More women are pursuing degrees in science, and these graduates constitute the youngest, more-recently trained scientists in their disciplines (Hornig, 1984). However, little is known about the professional characteristics and responsibilities of women scientists working in an academic setting. Most current research examines sex discrimination, sex role stereotypes, and sex biases (Butler & Marzone, 1980). However, Dresselhaus (1984), stresses the importance of studying the responsibilities of women scientists. Specifically, she studied the responsibilities of women faculty in engineering schools. At the same time, with regard to women in the agricultural sciences at universities, little descriptive information is available.

Purpose of the Study

This research effort was the first national study to focus specifically on women agricultural scientists in academic settings. The main purpose of the study was to characterize women scientists in colleges of agriculture at the 70 land grant universities in the United States. Specifically, the study of women agricultural scientists was designed to investigate the following research objectives:

(1) to describe their academic background and current positions;
(2) to describe their teaching, research, and service responsibilities; and
(3) to provide demographic data.