Helmers, J.E. 1993. Returns to Bachelor's and Master's Degrees in Agribusiness Economics: the Non-Land-Grant Experience. MS Thesis. Dept. of Plant and Soil Science, Southern Illinois University at Carbondale College of Agriculture. Carbondale, IL 62901.

Illinois Board of Higher Education (IBHE). 1997. Transfer and Articulation. Item #4. [Internet document], accessed July 1999; available from http://www.ibhe.state.il.us/articula.htm. May 6.

Koutsoyiannis, A. 1977. Theory of Econometrics. 2nd ed. Totowa, New Jersey: Barnes and Noble Books.

Rudd, R. D., R. A. Smick-Attisano, and T. S. Hoover. 1996. Identification and Analysis of Characteristics of Associate Degree Agricultural Students. NACTA Jour. 40(2): 32-34.

Steering Panel, Board of Higher Education, Illinois Community College Board. 1996. Articulation Recommendations for Baccalaureate Majors. Pub. Illinois Board of Higher Education (IBHE) Springfield, IL. July.

Relationships Between Student/Course Characteristics and Student Evaluations of Teaching Quality

Matt Baker[1], Rick Rudd[2], and Tracy Hoover[1], Department of Agricultural Educations University of Florida, Gainesville, FL 32611-0540

Abstract

A relational study was conducted to determine associations among student final grade, class size, course level (undergraduate or graduate), time of evaluation (fall or spring semester), type of course (applied behavioral science or biological/physical science) and student evaluations of teaching and course quality. The student evaluation instrument contained the following measures: (1) overall course quality, (2) instructor's ability, (3) overall instructor quality, (4) exams and assignments, and (5) instruction in the laboratory. The quantitative results of this study revealed that statistically significant positive bivariate correlations existed between class size and student perceptions of the instructor's ability, overall instructor quality, and exams and assignments. Positive but low significant relationships were also found between student course grade and these same teaching quality measures. Grades were also significantly positively correlated with perceived quality of laboratory. Students enrolled in graduate level courses were more pleased with the quality of instruction in three of the five measures. Overall, student evaluations were not drastically tainted by these factors.

[1]Associate Professor
[2]Assistant Professor

Introduction

Student evaluations of teaching quality provide essential information related to numerous decisions in higher education (Braskamp et al., 1983, 1984b). They provide information to students (in terms of course selection), colleagues (for tenure, promotion, merit pay, etc.), and the instructor (for self-improvement). Increasingly the validity of student evaluations and the quality of evidence that they provide to the personnel function have been questioned. Supporters of student evaluations perceive students as the most reliable evaluators because they are recipients of the instruction on a continuous basis. Frey (1976) stated: "because students are the only regular observers in college classrooms, reports about their classroom experiences provide unique information about the teacher and the teaching environment" (p. 327). Supporters view student evaluations as effectively providing summative evidence related to instructional quality, as well as providing formative evidence for instructional improvement. There are many critics, however, that are concerned about the use of these evaluations (Greenwald, 1997, Greenwald and Gillmore, 1997; Wilson, 1998). These critics feel that the classroom is transformed into a popularity contest. In their view the teacher's role is changed from that of educator to entertainer.

Braskamp et al. (1984a) advanced a model that contained four factors that could have a positive or

negative impact upon student evaluations of instruction. These factors were: (1) the nature of the course (required/ elective, course level, class size, and discipline), (2) the instructor (rank, sex of instructor, years teaching), (3) the student (expected grade, prior interest, major/ minor, sex and personality characteristics), and (4) the evaluation instrument (placement of items, number of response alternatives, negative wording of items, labeling all scale points vs. labeling only end points). This study examines the influence of selected student and course factors upon student evaluations. The specific factors included in this study related to the nature of the course were: (1) class size, (2) student level (undergraduate or graduate), (3) time of evaluation (fall or spring semester), and (4) type of course (applied behavioral science or biological/physical science). The single student factor included in the study was final grade.

In an extensive literature review Coburn (1984) conceded that researchers were divided about the impact of class size on teacher evaluations. Schlenker and McKinnon (1994) identified the ideal class to be from 16-30 people. However in a study of over 4000 classes, Wood et al. (1974) reported that a curvilinear relationship exists between student evaluations and class size. Studies on the effect of course level upon student evaluations are not decisive. In one study Gigliotti and Buchtel (1990) reported that there was an influence on student evaluations. However, their study was limited to undergraduate students only. Conversely, Stufflebeam (1988) found no significant differences on student evaluations based upon an upper division or lower division classification of undergraduates. However, Aleamoni (1980) points out that the higher the level of education of the student, the higher the evaluation.

Although the research related to the time that student evaluations are completed almost exclusively focuses upon the consistency that delayed evaluations would have upon student evaluations, Frey (1976) found very little correlation to the time in the semester the evaluation was given and the evaluation. Shapiro (1990) found "the timing of evaluations appears to have little effect, whether the evaluations are administered at the end, the next semester or even years later" (p. 137). However, the researchers for this study hypothesized that during the course of an academic year, students would tend to give fall semester courses higher evaluations than spring semester courses due to a basic fatigue factor. It also stands to reason that academics with heavy teaching loads would be fresher in the fall semester and more exhausted in the spring semester.

Studies indicate differences between student evaluations and the type of course. In his literature review, Shapiro (1990) reported that student evaluations tended to be lower in the sciences. Feldman (1978) analyzed 11 studies on student evaluations based upon academic fields. He reported that the social sciences tended to be "in the medium or low third of rankings" and that hard science courses "are also usually in the lower two thirds of the rankings" (p. 222). The results of the research on the impact of student final grades upon student evaluations are also mixed. Shapiro (1990) found that class grade accounted for six percent of the variance in student evaluations. Gigliotti and Butchtel (1990) on the other hand, stated that "there is very little evidence to support the popular beliefs that easy or hard grading affects evaluations of instructors" (p.350).

The purpose of this study was to determine the influence of selected student and course factors upon student evaluations. The specific objectives of the study were to: (1) describe courses in terms of academic unit, class size, level, final grade, and student evaluations; (2) identify relationships between student evaluations and class size and final grade; and (3) determine if student evaluations differed significantly based upon type of course and time of evaluation.

Materials and Methods

The population of this study consisted of final average student evaluations from 1,264 courses taught in the College of Agriculture at the University of Florida between Fall semester 1995 and Spring semester 1997 excluding summers. A random sample by semester was drawn from the total number of courses using the guidelines proposed by Krejcie and Morgan (1970), resulting in an overall sample size of 678. The evaluation instrument was developed by a College of Agriculture Committee, and was later adopted for university-wide use. The instrument utilized a five-point Likert-type scale (1=poor, 5=excellent) in measuring the following: (1) overall course quality (one item), (2) instructor's ability (two items), (3) overall instructor quality (two items), (4) exams and assignments (2 items), and (5) instruction in the laboratory (four items). The overall post-hoc reliability coefficient of the instrument was r=.62. The data were analyzed utilizing SPSS/7.5 for Windows software. Descriptive statistics, Pearson Product Moment correlations, and ANOVA were used to summarize and analyze the data.

Results and Discussion

Courses from 36 prefixes were represented in the sample. Food and Resource Economics represented the greatest percentage of the courses with 12.7%. Animal Science and Agricultural Education and Communication were next with 6.8% and 6.7% respectively. Enrollment in the courses varied tremendously. The largest course

enrollment was 653, and the smallest enrollment was three. The average class size was 29 (SD=62.30). Over 62% of the courses were offered at the upper division undergraduate level (junior or senior level). Six and one-half percent were lower division undergraduate courses, while almost one-third were graduate level courses. The average overall course grade was a 3.34 (SD=0.47) on a four point scale. In terms of overall student evaluations, Table 1 indicates that on average, students rated the courses and instructors very favorably. Little variability was observed on final mean rating score.

According to Table 2, statistically significant correlations were discovered between student evaluations and the number of students enrolled on three of the measures. In terms of magnitude, these correlations did not appear to be very meaningful. All but one of the constructs between student evaluations and average course grade were significant. Once again, the correlation with the largest magnitude was responsible for explaining less than eight percent of the variance (r=.28) in student evaluation. Grades and class size are not very strong in affecting student evaluations of course and instructor quality. Overall the student evaluations were statistically significant and positive toward the instructor, course structure (exams and assignments), and laboratory.

Table 1. Overall Course Evaluations.

Measure	Mean[z]	Standard Deviation
Overall course quality	4.23	1.53
Instructor's ability	4.30	0.44
Overall instructor quality	4.35	0.50
Exams and assignments	4.29	0.50
Instruction in the laboratory	4.28	0.58

[z] Based upon a five-point Likert type scale, 1=poor, 5=excellent

Table 2. Relationshipsy of Student Evaluations[z] by Enrollment and Grade

Measure	Enrollment (r)	Enrollment (r^2)	Grade (r)	Grade (r^2)
Overall course quality	.05	.003	.07	.005
Instructor's ability	.13*	.017	.28*	.078
Overall instructor quality	.11*	.012	.23*	.053
Exams and assignments	.11*	.012	.23*	.053
Instruction in the Laboratory	.06	.004	.21*	.044

y Pearson Product Moment correlations
z Based upon a five-point Likert type scale, 1=poor, 5=excellent
* implies significance at the .05 level

Positive low relationships were found between instructor's ability (r=.13), overall instructor quality (r=.11), exams and assignments (r=.11) and enrollment. The current study also supports the results of Wood et al. (1974) who found that instructors teaching larger classes tended to receive more favorable evaluations. Positive low relationships were discovered in the current study between instructor's ability (r=.28), overall instructor quality (r=.23), exams and assignments (r=.25), instruction in the laboratory (r=.21) and average course grade. This is supportive of the

findings by Shapiro (1990) who found that grades explained about six percent of the variance in student evaluations; in our study, the percentage of the variance explained by enrollment and grade ranged from a low of less than one percent to almost eight percent on the subscales of the student evaluation instrument.

When graduate and undergraduate students were separated in the correlational analysis, only one significant correlation was found between enrollment and graduate student evaluations. The significant correlation was between exams and assignments and class size (r=.14). None of the undergraduate student evaluation constructs were significantly related to class size. Relationships between grade and student evaluations of courses were also examined using this same dichotomy. Only one significant correlation was found between enrollment and graduate student evaluations. The significant correlation was between exams and assignments and instruction in the laboratory (r=.21). However, four of the five undergraduate student evaluation constructs and grades were significantly related. Undergraduate students receiving higher grades were more satisfied with the instructor's ability (r=.26), overall instructor quality (r=.21), course exams and assignments (r=.23), and laboratory instruction (r=.20).

According to Table 3, graduate level courses were rated significantly higher in terms of the instructor's ability, overall instructor quality, and exams and assignments, when compared to undergraduate level courses. It should be noted here that although a few instructors teach courses at both levels, this course dichotomy primarily examines two unique populations. This finding is in agreement with the comprehensive synthesis of research reported by Aleamoni (1980).

Table 3. ANOVA: Student Evaluations by Level

Measure	Undergraduate Mean	Graduate Mean	R2	F	p
Overall Course Quality	4.23[1]	4.25	<.001	0.05	.818
Instructor's ability	4.27	4.38	.012	8.47	.004
Overall instructor quality	4.32	4.42	.008	5.78	.017
Exams and assignments	4.26	4.36	.009	6.21	.013
Instruction in Laboratory	4.27	4.29	<.001	0.60	.806

[1] Based upon a five-point Likert type scale, 1=poor, 5=excellent

Table 4 reveals that no significant differences were found in student evaluations based upon the type of course (hard vs. soft science dichotomy). Behavioral science courses included Human Resource Development, Agricultural Education and Communication, and Food and Resource Economics. Similarly, no significant differences were discovered in student evaluations and semester that the course was taught (see Table 5). Thus the hypothesis proposed by the authors related to student fatigue and teacher enthusiasm impacting student evaluations was not supported by the data. These same analyses were conducted examining the graduate and undergraduate student samples. No significant classification effect (graduate vs. undergraduate) was found for student evaluations based upon type of course (hard science vs. soft science) or semester (fall vs. spring).

Table 4. ANOVA: Student Evaluations by Type of Course.

Measure	Applied Behavioral Science Mean	Biological or Physical Science Mean	R^2	F	P
Overall Course Quality	4.17[z]	4.25	<.001	0.268	.605
Instructor's ability	4.29	4.31	<.001	0.150	.699
Overall instructor quality	4.35	4.35	<.001	0.018	.894
Exams and assignments	4.29	4.29	<.001	0.018	.894
Instruction in Laboratory	4.19	4.29	0.004	2.18	.140

[z] Based upon a five-point Likert type scale, 1=poor, 5=excellent

Summary

In terms of the model advanced by Braskamp et al. (1984a), this study provides support of student grade in class as being related to higher student evaluation of instructor scores. However, the strongest relationship on the student evaluation measures was found to explain less than eight percent of the variance. As per the nature of the course, class size and student level (graduate or undergraduate) also were related to student evaluations, with instructors of larger classes or graduate courses receiving higher scores on evaluations.

Table 5. ANOVA: Student Evaluations by Semester Taught

Measure	Fall Semester Mean	Spring Semester Mean	R2	F	p
Overall Course Quality	4.29[z]	4.18	.001	0.845	.358
Instructor's ability	4.31	4.29	.001	0.412	.521
Overall instructor quality	4.36	4.34	<.001	0.301	.584
Exams and assignments	4.30	4.28	<.001	0.291	.590
Instruction in Laboratory	4.23	4.32	.006	2.88	.090

z Based upon a five-point Likert type scale, 1=poor, 5=excellent

Literature Cited

Aleamoni, L. M. 1980. The use of student evaluations in the improvement of instruction. NACTA Jour. 24(3): 18-21.

Braskamp, L., E. Brandenburg, E. Kohen, J. Ory, and P. Mayberry. 1983. Guidebook for evaluating teach ing, Part I. NACTA Jour. 27 (4): 29-33.

Braskamp, L., E. Brandenburg, E. Kohen, J. Ory, and P. Mayberry. 1984a. Guidebook for evaluating teaching, Part II. NACTA Jour. 28 (1): 19-24.

Braskamp, L., E. Brandenburg, E. Kohen, J. Ory, and P. Mayberry. 1984b. Guidebook for evaluating teaching, Part III. NACTA Jour. 28 (4): 27-31.

Coburn, L. 1984. Student evaluations of teacher performance. Princeton, NJ: Educational Testing Service.

Feldman, K. A. 1978. Course characteristics and college students' ratings of their teachers: What we know and what we don't. Research in Higher Education 9 (1): 199-142.

Frey, P. W. 1976. Validity of student instructional ratings: Does timing matter. Jour. of Higher Education 47 (3): 327-336.

Gigliotti, R. J. and F. S. Buchtel. 1990. Attributed bias and course evaluation. Jour. of Educational Psychology 82 (1): 341-351.

Greenwald, A. G. 1997. Validity concerns and usefulness of student ratings of instruction. American Psychologist 52 (11): 1182-1186.

Greenwald, A. G., and G.M. Gillmore. 1997. Grading leniency is a removable contaminant of student ratings. American Psychologist 52 (11): 1209-1217.

Krejcie, R.V. and D.W. Morgan. 1970. Determining sample size for research activities. Educational and Psychological Measurement 30: 607-610.

Schlenker, D. E. and N.C. McKinnon 1994. Assessing fac ulty performance using the student evaluation of instruction. Boston, MA: Atlantic Boston Col lege.

Shapiro, G. E. 1990. Effect of instructor and class character istics on student's class evaluation. Research in Higher Education 31 (6): 135-147.

Stufflebeam, C. E. 1988. Relationships between student characteristics and their ratings of professors. NACTA Jour. 32 (1): 7-10.

Wilson, R. 1998. New research casts doubts on value of student evaluations of professors. The Chronicle of Higher Education 54 (2): 12-16.

Wood, K., A. S. Linsky, and M.A. Straus. 1974. Class size and student evaluations of faculty. Jour. of Higher Education 45 (7): 524-534.