

Utah State University

DigitalCommons@USU

---

All Graduate Theses and Dissertations

Graduate Studies

---


5-1976

## Correlates of Course Ratings

Jan Krambule

*Utah State University*

Follow this and additional works at: <https://digitalcommons.usu.edu/etd>

 Part of the [Psychology Commons](#)

---

### Recommended Citation

Krambule, Jan, "Correlates of Course Ratings" (1976). *All Graduate Theses and Dissertations*. 5807.  
<https://digitalcommons.usu.edu/etd/5807>

This Thesis is brought to you for free and open access by the Graduate Studies at DigitalCommons@USU. It has been accepted for inclusion in All Graduate Theses and Dissertations by an authorized administrator of DigitalCommons@USU. For more information, please contact [digitalcommons@usu.edu](mailto:digitalcommons@usu.edu).



CORRELATES OF COURSE RATINGS

by

Jan Krambule

A thesis submitted in partial fulfillment  
of the requirements for the degree

of

MASTER OF SCIENCE

in

Psychology

Approved:

UTAH STATE UNIVERSITY  
Logan, Utah

1976

## ACKNOWLEDGMENTS

I wish to thank all those who assisted in the completion of this study.

To my committee chairman, Dr. Elwin Nielsen, I express sincere appreciation for his guidance, suggestions and encouragement before and during the writing of this thesis.

Appreciation is expressed to Dr. Keith Checketts, whose statistical knowledge and willingness to contribute time greatly aided in the completion of this research project.

To Dr. Bill Dobson, I express gratitude for his consistent support, encouragement and caring.

Also, I would like to give special mention to Dr. Rex L. Hurst for his help with some of the statistical aspects of this study.

Finally, I would like to thank my husband, Dar, for his encouragement and support.

Jan Krambule

## TABLE OF CONTENTS

	Page
ACKNOWLEDGMENTS . . . . .	ii
LIST OF TABLES . . . . .	iv
ABSTRACT . . . . .	v
NATURE AND BACKGROUND OF THE STUDY . . . . .	1
Background Information . . . . .	1
Need for the Study . . . . .	2
Statement of the Problem . . . . .	4
RELATED RESEARCH . . . . .	5
Studies Concerning Student Evaluation of Faculty . . . . .	5
What Students Look For in University Faculty . . . . .	6
Correlates of Course Ratings . . . . .	8
METHOD . . . . .	19
Purpose . . . . .	19
Procedures . . . . .	19
RESULTS . . . . .	24
Description of the Sample . . . . .	24
Regression Analysis . . . . .	24
Regression Analysis with Question 23 Removed . . . . .	29
Summary . . . . .	29
DISCUSSION . . . . .	32
Limitations . . . . .	33
Conclusions . . . . .	34
Recommendations . . . . .	34
LITERATURE CITED . . . . .	37
APPENDICES . . . . .	40
Appendix A. USU Faculty Evaluation Form . . . . .	41
Appendix B. Memo to USU Faculty . . . . .	44
Appendix C. Scatter Diagrams Comparing Total Mean Scores and the Selected Variables . . . . .	46



## LIST OF TABLES

Table	Page
1. Regression Analysis . . . . .	25
2. Means for Sample Size . . . . .	27
3. Means for Colleges . . . . .	27
4. Means for Course Levels . . . . .	28
5. Regression Analysis with Question 23 Removed . . . . .	30

## ABSTRACT

## Correlates of Course Ratings

by

Jan Krambule, Master of Science

Utah State University, 1976

Major Professor: Elwin Nielson  
Department: Psychology

The purpose of this study was to assess the extent to which selected variables are related to scores on the Utah State University Faculty Evaluation Questionnaire (USU FEQ). The variables examined were size of class, level of class, college under whose auspices a class is offered and whether the class is required or elective. In addition, the relationship between total mean score and responses to question 23, a percentage ranking of instructors, was assessed.

The fall quarter, 1975, FEQ results were used in this assessment. Employing the same procedures as had been followed in previous quarters, 315 courses were evaluated. Courses at all levels (100-700) and within all colleges of the University were sampled.

Through multiple regression analysis, it was found that 87% of the variance in total mean scores was explained by the selected variables. Question 23 was the outstanding contributor; therefore, the analysis was repeated with this variable removed. As a result, 21% of the variance was explained by the remaining variables.

Question 23 was highly related to total mean score. Reasonably accurate prediction of instructors ratings can be made from a knowledge

of responses to this question.

A small negative correlation was found between sample size and total mean score. Instructors of larger classes tended to rate lower than instructors of smaller classes. A small contribution to the explained variance was made by the variable of class size. This contribution has little practical significance.

Instructors of different colleges received different average ratings. Those colleges most closely related to high total mean scores were Education, Family Life and Humanities, Arts and Social Sciences. Those colleges receiving lowest ratings were Engineering and Science. Knowledge of the college under whose auspices a course exists improves prediction of instructors ratings on the USU FEQ.

Whether a course was required or elective and the level of a course had little relation to the ratings instructors received.

The outcomes of this investigation may be limited by any one or a combination of the following:

1. The results of this study were obtained from fall quarter evaluations. Ratings may relate to the quarter in which the course was evaluated.
2. The USU population was the only university used in the sampling. Results, therefore, will be generalizable to USU alone.
3. This study has been concerned with environmental variables, i.e., class size, level, college, required vs. elective. Student, class, or teacher characteristics have not been considered. Significant relationships may exist in these areas.

4. Within University policy, professors are allowed to choose the quarter in which they will be evaluated. Therefore, some volunteer effects may be present in this study. However, extra encouragement to evaluate fall quarter was given to teachers to help control for these effects.

From this study of the relationship between selected variables and total mean scores, the following may be concluded:

1. Question 23 could be used to obtain a quick, easy estimate of student ratings of an instructor.

2. The significant differences between ratings of instructors in different colleges receive presents an area of concern. When comparing instructors from different colleges, precautions may need to be taken. Differential norms for the colleges could be considered.

3. The size of an instructor's class is of no significant concern when interpreting FEQ results. The results of this study show no basis for the employment of differential norms for differing class sizes.

4. No basis for considering whether a course is required or elective and level of a course when interpreting questionnaire results is given.

From the above summary and conclusions the following recommendations are made:

1. The college under whose auspices an instructor teaches should be considered when interpreting FEQ results. College norm groups should be considered for inclusion on results printouts.

2. The significantly higher ratings received by instructors

in the College of Education should be investigated. The following questions arise:

- A. Do students rate instructors higher if instructors give them higher grades? (Grades received in Education are higher.)
- B. Are the students in the college more empathetic to their instructors and, therefore, more lenient in their ratings?
- C. Is the course content less demanding or more interesting so as to receive higher student ratings?
- D. Are, in fact, the instructors in the College of Education better teachers? If so, what are the characteristics that make them better?

These questions and others should be researched.

3. Since relationships may vary according to the particular aspect of teaching performance that the student is asked to rate (Clark & Keller, 1954), a factor analytic study including the selected variables of this study and all of the questions on the USU FEQ may be profitable.

4. Scores on the USU FEQ may be related to the department in which the course is taught. Correlations might be determined for the departments which have a large enough N.

5. Determination of whether students' subjective criteria in rating faculty match the faculty members' goals in teaching may disclose some valuable information about the USU FEQ.

6. The benefits from student evaluation of instruction can only exist to the extent that ratings represent valid appraisals of classroom

instruction. Research must be conducted in answer to the question,  
"Is the USU Faculty Evaluation Questionnaire a valid instrument?"

(60 pages)

## Nature and Background of the Study

### Background Information

The Utah State University Faculty Evaluation Questionnaire (USU FEQ) is an instrument used to collect student assessments of aspects of teaching in college courses. The results obtained with the questionnaire serve three major functions. The results are available to students making class selections and, at this time, formal publication of questionnaire results is being considered by the Associated Students of USU. In addition to student use, the questionnaire is used by faculty as a source of information about student reactions to their teaching. The results of the student ratings are also used by department heads, deans, and the central administration in their considerations of staff members for promotion, salary adjustments, tenure, and, in some cases, in determining whether a person will be reemployed.

The rating instrument initially used by USU was the Illinois Course Evaluation Questionnaire. At the University of Illinois, each instructor's results are compared to appropriate University of Illinois campus norm groups. Comparisons are made with other instructors of his or her own academic rank, with those teaching at a similar course level (100, 200, etc.), as well as with the all-university population (Spencer, 1968). The Illinois form was discontinued at USU in 1974-75 and a new form was implemented (Appendix A). At Utah State University, the results for each instructor are compared with the all-university norms only.

USU Faculty Evaluation Questionnaire results are compiled for each class on a computer printout. Responses to each of the 24 items are reported in summary form as mean, standard deviation, and norm decile based on the population of USU professors who have been rated by students using the scale since 1973. The mean, standard deviation, and decile are also reported for an overall or total score. This is obtained by summing responses to all items, with the exception of Item 16, i.e. "I was interested in the course when I registered for it." In addition, if a large percent of the responses were in the "Not Applicable" category, the item is not used in computing total scores.

#### Need for the Study

Since student evaluation of instruction is widely used at USU, it might be expected that research pertaining to the rating instrument would have been done. In particular, research might have been done which indicates that the total mean score provides valid information about aspects of instructors' teaching. No such research has been reported, however.

Use of faculty evaluation results without research concerning the evaluation method is not limited to USU. To assess the current status of student evaluation of instruction, Bejar (1973) made a survey by mail of 333 American university administrators responsible for faculty evaluation. A 68% return was obtained. Bejar concluded that although there has been an increase in the popularity of student ratings as a means of evaluating faculty performance, as well as an increase in the frequency with which evaluation results are used in decisions concerning faculty promotion, salary adjustments, tenure, and reemployment,



there has been minimal research validating the student ratings.

Without research indicating that the rating instrument is a valid indicator of student assessment of teaching, use of such data in decision making hardly seems justified. Faculty evaluations might even adversely affect the quality of education. Recognition of the negative potential of student evaluations of faculty and consideration of the broad effects on quality of education and educational institutions was stressed by Zelby (1974). For example:

Indiscriminate use of SFE [Student/Faculty Evaluation] will increase the gap between first-rate and second-rate institutions--first-rate institutions will continue to attract more demanding students, a fact that will be reflected in SFE's where as second-rate institutions, in an effort to maintain enrollment may tend to formulate SFE's that emphasize popularity and mediocrity of education. (p. 1270)

There may be long-range deleterious results unless more careful consideration is made of faculty evaluations before they are used extensively. (p. 1267)

Comparison, or relative standing with other instructors, is important in interpreting faculty evaluation results. There is a question, however, as to whether the total mean score taken alone is a valid indicator of teaching effectiveness. Is this statistic positively related only to quality of teaching performance, or is it also related to variables such as the size of an instructor's class?

Eble (1970), in reviewing Hildebrand's and Wilson's study of variables related to student ratings of instruction at the Davis Campus of the University of California, said:

In general, student ratings of best teachers showed only negligible correlations with academic rank of instructor, class level, number of courses previously taken in the same department, class size, required versus optional course, course in the major of respondent, class level of respondent, grade point average, and expected grade in the course.  
(p. 91)

In contrast, other research has indicated that upper division students tend to give more favorable ratings than did lower division students (Lovell & Haner, 1955), large classes receive lower ratings than small ones (McDaniel & Feldhusen, 1970), and required courses are rated lower than elective courses (Cohen & Humphreys, 1960). In addition, significant positive relationships have been found between students' grades and their ratings of instructors and courses (Treffinger & Feldhusen, 1970).

The research is contradictory. Mean scores may be strongly related to class size, level of class, college under whose auspices a course is taught, and whether the class is required or elective. If such relationships exist, these variables should be considered when interpreting instructors' questionnaire results.

#### Statement of the Problem

The problem is, then, the lack of information about the extent to which class size, class level, college under whose auspices a class is offered, and required versus elective courses, are related to instructors' mean scores on the USU Faculty Evaluation Questionnaire.

## Related Research

### Studies Concerning Student

#### Evaluation of Faculty

The past several years have seen a striking increase in attempts to formally evaluate college teaching. College faculties and administrators have been giving increased attention to students' opinions of their courses and instructors (U.S. Presidents Commission, 1970). As early as 1961, Gustad's survey into the methods of teacher evaluation used by 584 colleges and universities revealed that formal student appraisals were cited most often.

The widespread use of student evaluation of college faculty has led to a multitude of research studies, as seen in the literature. These studies have attempted to answer several questions. The basic questions are: "Are students capable of evaluating instructors?", "Are student's ratings reliable, valid and useful?", and "What are the explanatory variables, or correlates, which provide prediction of general course ratings?"

Specific questions that studies have attempted to answer follow. What are students' criteria of effective teaching? Is an instructor's ability to entertain correlated with high student ratings? Do students' ratings match supervisors' and colleagues' ratings? What is the relationship between student ratings and student achievement? How does student-teacher interaction, student's sex, major and college year affect student ratings? What are students' opinions concerning the value of student ratings? To what extent does an instructor's research

production correlate with student ratings of effective teaching? What, if any, instructor personality traits correlate with high student ratings? How should student rating instruments be developed?

The findings of most of these studies are not directly relevant to the problem, "the lack of information about the extent to which class size, class level, class college and required versus elective courses, are related to instructors' mean scores on the USU Faculty Evaluation Questionnaire." But, in answer to the basic questions, Costin, Greenough, and Menges (1971), in summarizing an extensive review of literature, say this:

A review of empirical studies indicated that student ratings can provide reliable and valid information about the quality of courses and instruction. Such information can be of use to academic departments in constructing normative data for the evaluation of teaching and may aid the individual instructor improving his teaching effectiveness. (p. 530)

#### What Students Look For in

#### University Faculty

Before reviewing current research concerning the correlates of student ratings, specifically, class size, class college, class level and required versus elective course, a further analysis of what characteristics students look for when rating college faculty is needed.

In determining what particular characteristics of college teachers, as perceived by students, were related to students' overall opinions of their teachers' effectiveness, French (1957) found that the 10 items which contributed most to student overall judgment at the University of Washington were: (a) interprets abstract ideas and theories clearly, (b) gets students interested in the subject, (c) has increased

my skills in thinking, (d) has helped broaden my interests, (e) stresses important material, (f) makes good use of examples and illustrations, (g) motivated me to do my best work, (h) inspires class confidence in his knowledge of the subject, (i) has given me new viewpoints or appreciations, and (j) is clear and understandable in his explanations.

Coffman (1954), as a result of research at Oklahoma Agriculture and Mining College, rank ordered four factors of "effectiveness." Factor A was named "empathy," which included ability to arouse interest, humor, interpersonal relations, and tolerance. Factor B was identified as "organization." Factor C represented the teacher's personality; and his verbal fluency ranked last as Factor D.

More recently, students were asked to describe the most effective college teacher they had ever had (Crawford & Bradshaw, 1968). The four most frequently mentioned characteristics were: (a) thorough knowledge of subject matter, (b) well planned and organized lectures, (c) enthusiastic, energetic, lively interest in teaching, and (d) student-oriented, friendly, willing to help students.

The most important criteria in describing an "ideal" professor (in rank order of importance) as determined by Gadzella (1968) were: (a) knowledge of subject (subject mastery), (b) interest in subject (enthusiasm), (c) flexibility (ability to meet student needs), (d) daily and course preparations (well organized), and (e) vocabulary (ability to explain clearly). The least important characteristics were: (a) the professor as a writer, (b) as a participator in the community, and (c) as a researcher.

Characteristics mentioned by other authors follow. Eble (1970) lists competence, caring, energy, imagination and sense of proportion. Musella and Rusch (1968) found, among the most frequently mentioned characteristics, expert knowledge of subject matter, systematic organization of course content, ability to explain clearly, enthusiastic attitude toward the subject, and ability to encourage thought. Knowledge of subject matter, interest in the subject, being well prepared for class, and motivating students to do their best were listed by Downie (1952). Smith (1944) mentioned sympathetic interest in students, knowledge of subject matter, ability to stimulate intellectual imagination, and ability to organize and put across subject matter.

As is indicated, the results of student ratings have produced many characteristics for "effective" instructors. However, the investigations are not at variance and, although phrased somewhat differently, provide reasonably consistent findings.

#### Correlates of Course Ratings

The studies of correlates of course ratings have been concerned with the prediction of students' ratings of faculty, i.e., investigators have tested the strength of relationships between ratings and a number of variables having potential explanatory power. Among these explanatory variables are those environmental factors, which can not be controlled by the instructor, and which may influence an instructor's rating. Delineated in the problem of this thesis are four of these factors. The findings about the relationships of these variables to student rating scores are contradictory. To illustrate, a review of the literature on each of these variables follows.

Class size. The prediction that teachers of large classes will receive lower ratings, perhaps because students generally prefer small classes which permit more student-teacher interaction, has been supported by the results of three major studies. Kohlan (1973) administered instructor evaluation questionnaires to 271 undergraduates in eight arts and science, business administration and education classes at a midwestern university. The Instructor Evaluation Questionnaire was administered after the second class hour and again during the last week of the semester. Following an analysis of variance of mean ratings, he concluded, among other things, that students in small classes tended to evaluate instructors more positively than did students in large classes.

Data which support Kohlan's conclusions are given by McDaniel and Feldhusen (1970) in a study which was conducted at Purdue. A slightly modified version of the Purdue Course Evaluation Questionnaire was used with a sample of 76 university professors. A total of 4,484 students completed rating scales in classes ranging in size from 9 to 408 students. Multiple correlations were computed for class size (and some other variables) as independent variables with composite ratings for course and instructor as dependent variables. Negative correlations were found between class size and ratings indicating that as class size increased, ratings decreased. According to the authors, "the results clearly indicate that the larger the class, the lower are ratings of instructional effectiveness. Perhaps the larger the class, the less able is he (the teacher) to secure feedback about individual students' needs to guide his teaching " (McDaniel and Feldhusen, 1970, p. 620).



In earlier research, Lovell and Haner (1955) also discovered a negative relationship between class size and high course ratings. They used a forced-choice type scale at Grinnell College which had been constructed by A. R. Rustebalke, a senior student in 1949. In computing  $t$  values for differences in mean ratings of instructors in various sized classes, they found that classes with sizes of 31 and up received significantly lower ratings than did smaller classes.

Other investigators have found that students assign higher ratings of overall teaching ability to instructors of large classes. However, these studies, as a rule, are less recent than those finding inverse relationships between class size and course ratings. Guthrie (1954), in a progress report on the evaluation of teaching at the University of Washington, states that students were not reluctant to assign high ratings of overall teaching ability to instructors of large classes.

Heilman and Armentrout (1936) report the results of the spring quarter (1935) faculty evaluation at Colorado State College. Faculty with a class size of 25 or more were asked to administer the Purdue Rating Scale for Instructors. Ratings of 46 teachers were made by 50 classes. Four classes were rated with a class size below 25. The range of class sizes was 17-121 with an average of 42. The product-moment coefficient for class size and the individual scale-means was found to be positive, .236. According to these results it is not certain that the size of the class making the rating has an influence upon the quality of the ratings, but the chances are slightly in favor of an increase of the rating with the size of the class.



Other studies indicate that there is little correlation between class size and course ratings. Hildebrand and Wilson's research at the Davis Campus of the University of California, as summarized in Eble (1970), support this finding. Of four surveys which were conducted, one included descriptions by 1015 students of the teaching of instructors previously rated by other students and/or faculty as to effectiveness of teaching. As a result of this survey (documented in the full report) only negligible correlations between students' ratings and class size were found.

No significant relationship between size of class and students' responses to the question: "Considering everything, how would you evaluate the instructor in this course?" was found by Solomon (1966). Two hundred and twenty-nine teachers from adult evening courses at five schools were surveyed and chi-square tests were applied to the results. This study may be limited as far as generalizability because of the relatively small size of all the classes. Class size was dichotomized at the approximate median which resulted in "large" classes having 10 or more students and "small" classes having 9 or less. Another limitation is that students of adult evening courses may not be representative of college students in general.

An extensive teacher-rating project was conducted at Brooklyn College in 1947 (Goodhart, 1948). Over 90% of the students rated each of five of their teachers using a questionnaire and rating form which had been pretested at Rutgers. It was found that teaching quality bears slight relationship, as far as student judgments go, to the size of class. Small classes did not receive more favorable results than

large ones. Goodhartz stated, "Where differences do occur, they are erratic and fail to emerge as a clear pattern."

Gage (1961) cited the results of analysis of student ratings of college teachers conducted in the College of Education at the University of Illinois. He reported a curvilinear relationship between ratings and class size: teachers in courses with 30-39 students received lower ratings than teachers in courses with either more or fewer students.

There is also evidence that the relationship between class size and student ratings of teaching may vary according to the particular aspect of teaching performance that the student is asked to rate. For example, Clark and Keller (1954) found a positive relationship between class size and students' opinions of how well the instructor was prepared for class, but a negative relationship between class size and the amount of original thinking students thought was demanded of them. In the same study, small classes were perceived by students as having greater feelings of good will between student and teacher, but larger classes were rated as making more effective use of class time.

Another example of this phenomenon was seen by Downie (1952). In the large classes (over 30), the instructional procedures, tests and quizzes, and value of the course received less favorable ratings. In the small classes, instructor-student relations received the harsher ratings. Included here were four items concerned with background of the instructor, knowledge of subject matter, impartiality, and the use of sarcasm and ridicule.

Class level. Studies concerning student evaluation of faculty have occasionally dealt with the relationship between class level (100, 200, etc.) and scores on faculty evaluation. Approximations to studies of class level are those concerned with the relationship between class or age of the student, i.e., graduates versus undergraduates, upper classmen versus lower classmen. Both types of investigations require some examination.

In his previously mentioned study at the University of Illinois, Gage (1953) found that teachers of lower-level courses consistently received less favorable mean ratings than did those of more advanced courses. At the University of Illinois (Spencer, 1968) comparisons of Faculty Evaluation Questionnaire results are made with instructors at similar course levels. This practice may suggest that the Illinois administrators recognize the differences that Gage found earlier. However, at Princeton (Eble, 1970) the practice of issuing percentage rankings of individual courses in relation to all other courses at the same level has been discontinued.

In 1947 a committee on Improvement of Instruction was appointed from the College of Science, Literature and the Arts at the University of Minnesota (Clark & Keller, 1954). This led to the development of an Instructor Rating Scale. Winter quarter 1949, 380 classes and a total of 15,000 students were surveyed with the scale. It was discovered that seniors and graduates were more favorable toward their instructors than lower classmen. This conclusion was based on classifications of these ratings by years in college (freshman-graduate) and by course numbers (0-49, 50-99, 100-199).

Kent (1967), in reviewing past research, noted that one factor significantly related to student ratings was the rater's class, in that graduates gave higher ratings than undergraduates. This finding is supported by Kohlan (1973). Kohlan, as mentioned previously, surveyed undergraduates in a variety of classes at a midwestern university. Through an analysis of variance it was discovered that upper classmen tended to evaluate their instructors more positively than lower classmen.

In 1949, the staff of the State College of Washington were asked, on a volunteer basis, to distribute faculty evaluation sheets to one of their classes. Sixteen thousand evaluation forms were completed. Using samples of 300, the ratings were studied by comparing percentages of some variables, including class of student. Upper-division students rated a few of the 36 items more favorably than lower-division students. These were: (a) opportunity to discuss tests, (b) introduction of new books and authors, (c) influence of the course on students taking another course in the same area, (d) value of laboratories, (e) extent of intellectual curiosity aroused.

The 1946 Indiana Conference on Higher Education conducted a study using the Purdue Rating Scale for Instructors at 14 institutions (Remmers & Elliot, 1949). In analyzing the results it was found that freshmen rated their instructors no higher or lower than did seniors. However, graduate students rated instructors higher than did undergraduates. Subjects for this study were volunteers and, although their confidentiality was insured, it is doubtful that these professors are representative of faculty in general.

In the Hildebrand and Wilson study at the Davis Campus of the University of California (Eble, 1970), student ratings of best teachers showed only negligible correlations with class level and class of respondent. Support for these results were given by Granzin and Painter (1973). They suggested that one might expect more advanced (or older) students to be more favorably disposed toward their courses. However, their study showed no support for this suggestion. In this study of 637 students in 17 courses offered in 11 departments at the University of Utah, a fairly high agreement ( $r = .73$ ) was found between undergraduate and graduate students' ratings of the same instructors.

In the 1966 investigation by Steward and Malpass at the University of South Florida, all faculty members were asked to administer the Course and Instructors Information Form to at least one class. A return of 1975 forms from 67 instructors teaching 54 courses was obtained. Chi-square analysis of the standard questionnaire was completed. No significant relationship between the class of the student and the ratings given to an instructor were found.

Class college. Few of the studies examined dealt with the relationship between the college classification of a course (Education, Science, etc.) and scores on faculty evaluation questionnaires. The Minnesota study, as reported by Clark and Keller (1954), did ask the question, "Are teachers in same fields rated higher than those in others?" After computing differences among groups, it was found that instructors in humanities and natural sciences seemed to be more favorably regarded than those in social studies at their university.

There is additional, indirect support for the hypothesis that, on the average, different colleges may receive different ratings. Spencer (1968) noted that at the University of Illinois, Faculty Evaluation Questionnaire results for each instructor are compared with college norms. At Princeton (Eble, 1970), it was suggested for the purposes of comparison, that averages be provided for all undergraduate courses in the same department. It was further suggested that other comparisons, i.e., same level and all university, be eliminated. An assumption underlying these practices may be that instructors in different subject areas do receive, on the average, different ratings. The college, under whose auspices a class exists, may have some power in predicting ratings faculty will receive.

In slight opposition to this assertion, Heilman and Armentrout (1936) state that no reliable differences were found between the ratings of the seven different divisions studied. However, it should be noted that these divisions were within only one college, the Colorado State College of Education.

Required versus elective courses. A number of studies have dealt with the relationship between required versus elective courses and scores on faculty evaluation questionnaires. Among the most recent of these was the Granzin and Painter study at the University of Utah which has been previously discussed. The correlation between ratings and required versus elective course were significant at the .001 level. Required courses fell lower in the ratings and electives drew significantly higher ratings.

Gage (1961) at the University of Illinois found that teachers of elective courses received consistently more favorable ratings than did instructors of required courses. He states that the difference was not only statistically but also substantively significant. Gage (p. 18) asks, "Do teachers operate more effectively in elective courses, or are the students in such courses easier to please, perhaps because they are better motivated?"

In a memorandum to the faculty in the University of Illinois Department of Psychology, Cohen and Humphreys (1960) noted that students required to take a course tended to rate it lower than did students who selected the course. Lovell and Haner (1955) at Grinnell College support this note. In comparing ratings in required and not required courses a "t" value of 6.26 was reported. The difference was significant with elective courses rating higher.

Consistent with their previously mentioned results, Hildebrand and Wilson (Eble, 1970) found only negligible correlations between student rankings of best teachers and required versus optional courses. Another study asserting negligible difference was Goodhartz (1948) at Brooklyn College. No conclusive evidence was reported for believing that the ratings given to an instructor are affected by the fact that the course he or she teaches is an elective or prescribed course.

At the State College of Washington, Downie (1952) reported only slight differences between the ratings of students in required and elective courses. Ratings of those in elective courses were higher on organization and presentation of material, laboratory work, and adequacy of the job performed by the instructor. Students in required



courses noted greater opportunity to discuss tests and a greater willingness of the instructor to give personal help.

The research data specific to the problem, "the lack of information about the extent to which class size, class level, class college, and required versus elective courses, are related to instructor's mean scores on the USU Faculty Evaluation Questionnaire" is limited. Studies that cite the above mentioned variables lack informative material concerning the statistical analysis and procedures employed. In addition, these variables were never the main concerns of the investigators, i.e., they were viewed as minor parts of the total study.

In reviewing the research, the question arises as to the generalizability of results from one university to another. Differences in sample size and general characteristics of students attending different universities lead to doubt in this area. It is highly probable, for example, that students at Princeton differ significantly in needs and motivations from students at USU. Also, the formats of the respective Student/Faculty Evaluations, which were not reported, may have been quite different.

Furthermore, as can be seen by the studies which have been cited, the research is contradictory. Scores on faculty evaluation questionnaires may be related to class size, level of class, college of the class and whether the class is required or elective. If such relationships exist, these variables should be considered when interpreting instructors' questionnaire results.



## Method

### Purpose

The purpose of this study was to assess the extent to which selected variables are related to scores on the USU Faculty Evaluation Questionnaire. These variables are size of class, level of class, college under whose auspices the course is offered and whether the course is required or elective. It is felt that through preliminary studies of this sort the question of the validity of the USU Faculty Evaluation Questionnaire eventually may be resolved.

### Procedures

Population and sample. The target population for this study was the professors whose students have used and/or will use the USU Faculty Evaluation Questionnaire. The sample used was those professors who had classes evaluated fall quarter, 1975. A memo (Appendix B) was sent to each department reminding faculty members of the location of the forms and the University's current policy and procedure concerning student/faculty evaluation. The following statement of policy was included in the memo:

Over the period of three years, a faculty member will be evaluated in every undergraduate course taught that has a class size of ten or more. Each faculty member teaching undergraduate classes must be evaluated in at least one course every year. The faculty member may choose the course to be evaluated. Faculty members may have as many courses as they wish evaluated. Course evaluation will not be administered before the sixth week of the quarter and should not be administered after the ninth week of the quarter.

As has been done previously, faculty members were asked to request forms from the Counseling and Testing Center. The instructors were asked to start the administration of the form and then to turn it over to a class member or secretary to return it to Counseling and Testing. The forms were processed by computer and printouts of results were returned to instructors after the grades for the class being evaluated were turned in. Printouts were also sent to department heads, college deans and to the student files.

Data and instrumentation. The data for this study were obtained from the computer printouts of faculty evaluation results for 315 courses evaluated fall quarter, 1975. The information collected from the printouts and coded on computer cards is as follows.

1. The overall mean score as indicated on page 3 of each printout. (This is the dependent variable with the variables listed below as independent variables.)
2. The class size as indicated by the number labeled "sample size" on the printout.
3. The class level as indicated by the course number on the identification portion of each printout. Course level is indicative of a course's level of difficulty, i.e., 100 level being introductory, 700 level being advanced graduate.
4. The college under whose auspices the course is offered, as indicated by the identification portion of each printout. This was coded on the computer cards by assigning a number to each college. These are: Agriculture, 1; Business, 2; Education, 3; Engineering, 4; Family Life, 5; Humanities, Arts and Social Sciences, 6; Natural Resources, 7; Science, 8.

5. Whether the course is required as indicated by the percentage of students reporting that the course was required. The alternatives for this question, listed on the FEQ, are, "This course is: A. required in my major, B. required in my minor, C. a part of my general education, D. a free elective." If the percentage for A+B+C was greater than 50%, a "1" was recorded on the computer card indicating that the course was required. If the percentage for D was equal to or greater than 50%, a "0" was recorded indicating that the course was elective.

6. In addition to the original problem statement, another variable for comparison to the mean was investigated. This was question #23, i.e., "Compared to other instructors I have had, this instructor is in the ... A. Top 10%, B. Top 25%, C. Middle, D. Bottom 25%, E. Bottom 10%." Each alternative was assigned a number, i.e., A-4, B-3, C-2, D-1, E-0, and an average similar to grade point average was computed and recorded on the computer cards. The purpose of this addition was to assess whether answering this simple question would obtain the same results as the entire questionnaire.

The instrument, the USU Faculty Evaluation Questionnaire, is composed of 24 computer scorable questions and 7 essay questions on the back. For the purposes of this study, the questions scorable by the computer were utilized. Sixteen of these questions are positively stated, i.e., "Class time was well used." These offer "not applicable," "strongly agree," "agree," "disagree" and "strongly disagree" as options for response. The next six questions ask the students to

rate textbooks, lectures, etc., with the options of "not applicable," "excellent," "good," "fair" and "poor" for response. The two remaining questions are: "This course is ... A. required in my major, B. required in my minor, C. part of my general education, D. a free elective," and, "Compared to other instructors I have had, this instructor is in the ... A. Top 10%, B. Top 25%, C. Middle, D. Bottom 25%, E. Bottom 10%."

The USU Faculty Evaluation Questionnaire, as it now stands, was developed by the University Faculty Evaluation Committee in 1973. This committee studied faculty evaluation forms from other universities and extracted ideas from them that seemed applicable to USU. Essentially, the form is still under development; therefore, adequate reliability and validity information is not available.

Statistical design. The original statistical method which was to be applied to the data of this study was a step-wise multiple regression procedure. This procedure would allow the determination of the relative contribution of each independent variable to variance in faculty evaluation scores. Each variable was to be eliminated, through computations, in reverse order of its contribution, that is, the one contributing the smallest amount was to be eliminated first.

A modification of this original method was made because of the contribution of question 23 to variance in faculty evaluation scores. Through a regression analysis, it was found that 87% of the variance was explained by the combination of all selected variables. Eighty-one percent of the variance in total mean score was explained by

question 23 alone. The remaining variables have a unique contribution of only 6%. Through stepping these variables out, beginning with the one that contributes least, little gain in information would be made.

As a result, the relationship between the independent variables and faculty evaluation scores were determined. Regression analyses were computed twice: once to assess the combined contribution of all of the variables to variance in total mean scores and once to assess the contribution of the variables excluding question 23.

## Results

### Description of the Sample

The sample consisted of all courses which were evaluated fall quarter of 1975 on the USU campus. A total of 8,277 forms were returned for 315 courses. These courses ranged in size from 3 to 268 with a mean class size of 26.3. Included were 90% required and 10% elective courses and courses at all levels (85 courses at 100 level, 49 at 200, 72 at 300, 36 at 400, 45 at 500, 24 at 600, 4 at 700). Courses from all colleges of the University (Agriculture, 10 courses; Business, 71 courses; Education, 51 courses; Engineering, 48 courses; Family Life, 16 courses; Humanities, Arts and Social Sciences, 41 courses; Natural Resources, 12 courses; Science, 66 courses) were also included. Total mean scores ranged from 2.5-4.0 with a mean of 3.2 (on a scale from 1.0-4.0).

### Regression Analysis

Table 1 illustrates the contribution of the independent variables to variance in total mean scores. Eighty-seven percent of the variance was explained by the combination of the independent variables. The most significant contribution was made by question 23.<sup>1</sup> A correlation coefficient obtained for question 23 in relation to total mean score suggested that 81% of the variance was accounted for by this variable.

---

<sup>1</sup>Compared to the other instructors I have had, this instructor is in the ... A. Top 10%, B. Top 25%, C. Middle, D. Bottom 25%, E. Bottom 10%.

Table 1  
Regression Analysis

Variable	DF	Mean Square	F Ratio	Coefficient
Required vs. Elective	1	.0213	1.9198	-.0244
Sample Size	1	.6088	5.4858*	-.0006
Question 23	1	16.3185	1470.3570**	.4220
<u>College</u>	7	.1020	9.1983**	
Agriculture	1	.0017	.1558	-.0119
Business	1	.0310	2.7935	.0224
Education	1	.2600	23.4259**	.0758
Engineering	1	.2105	18.9632**	-.0675
Family Life	1	.0567	5.1100*	.0552
HASS	1	.0009	.0776	.0047
Natural Resources	1	.0085	.7658	-.0252
Science				-.0535
<u>Level</u>	6	.0277	2.4945	
100 Level	1	.0738	6.6480	.0361
200 Level	1	.0092	.8318	-.1505
300 Level	1	.0149	1.3440	-.1671
400 Level	1	.0053	.4809	-.1229
500 Level	1	.0032	.2906	-.8981
600 Level	1	.0170	1.5329	-.2580
700 Level				1.2356
Error	298	.0110	RSQ = .8661	

\* Significant at 5% level.

\*\* Significant at 1% level.

Because of the dominance of contribution of question 23, a regression analysis was computed with the question removed. Without the question, 21% of the variance was accounted for. This analysis is discussed later in this chapter.

From a knowledge of responses to question 23, a reasonably accurate prediction of total mean scores can be made. For comparison purposes, question 23 could be used to get a quick, easy estimate of student evaluation of instructors.

As shown in Table 1, whether courses were required or elected had little relation to total mean scores. Contribution to the total explained variance was insignificant.

The contribution of sample size to explained variance in total mean score was significant at the 5% level. Knowledge of a class's size contributed slightly to the explanation of variance in total mean scores. The relationship was inverse, i.e., as sample size increased, mean scores tended to decrease. However, the contribution of sample size has little practical significance. Examination of the differences in means (Table 2) and a scatter diagram (Appendix C-4) illustrates this point. The most differing mean score was reported for classes with a size of 9 or less. Results are unreliable when obtained from groups with sizes less than 10.

The college under whose auspices a course was taught contributed significantly to the explanation of variance in total mean scores. Table 3 and a scatter diagram (Appendix C-3) illustrate the differences between means for the various colleges. Table 1 contrasts scores obtained by the colleges and the coefficients indicate whether a



Table 2  
Means for Sample Size

Sample Size	Mean	Sample Size	Mean
0-9	3.3	40-49	3.1
10-19	3.2	50-59	3.1
20-29	3.2	60-over	3.1
30-39	3.1		

Table 3  
Means for Colleges

College	Mean	College	Mean
Agriculture	3.0	Family Life	3.3
Business	3.2	HASS	3.3
Education	3.4	Natural Resources	3.0
Engineering	3.1	Science	3.0

college received scores above or below average for colleges. The coefficient for Science was obtained by computing the negative sum of the other college coefficients. The college coefficients were added and a negative sign put in front of the sum.

The colleges of Education and Family Life were significantly above average as compared to the other colleges. Therefore, it would be expected that instructors in these colleges would receive higher ratings. In contrast, the colleges of Engineering and probably, Science were significantly below average for colleges. Lower ratings for instructors within these colleges would be expected.

Course level contributes little to explanation of variance in total mean scores (Table 1). The coefficient for 700 level was obtained by computing the negative sum of the coefficients for the other levels. The 700 level may be a significant contributor; however, the sample size for this level was only 4. Any interpretation with this small a sample size would be questionable. Table 4 reports the means for each course level. Little difference is indicated.

Table 4  
Means for Course Levels

Levels	Means	Levels	Means
100 Level	3.2	500 Level	3.2
200 Level	3.2	600 Level	3.3
300 Level	3.1	700 Level	3.5
400 Level	3.2		

### Regression Analysis with Question 23

#### Removed

Table 5 illustrates the regression analysis without the effects of question 23. Twenty-one percent of the variance in FEQ scores was explained by the combination of the remaining variables. Through examination of F-ratios, it can be seen that the major contribution to the explained variance was made by colleges.

The colleges of Education, Humanities, Arts and Social Sciences, Engineering and Science contributed most to explaining the variance in total scores.

Sample size was the only other significant contributor to explanation of variance in total mean scores. Knowledge of sample size improves prediction of total mean scores only slightly, however.

Required vs. elective courses and level of the course contributed very little to the explained variance in total mean scores.

#### Summary

Eighty-seven percent of the variance in total mean scores was explained by the combination of the selected variables. Without question 23, 21% of the variance was explained.

Question 23 can be used to assess student evaluation of instruction with confidence that scores comparable to the total mean score will be obtained.

The size of an instructor's class and the rating received show a slight relationship. Lower ratings tend to be given in larger classes. This finding has little practical significance, however.

Table 5  
Regression Analysis with Question 23 Removed

Variable	DF	Mean Square	F Ratio	Coefficient
Required vs. Elective	1	.0189	.2883	-.0230
Sample Size	1	.2834	4.3118*	-.0013
<u>College</u>	7	.5006	7.6261**	
Agriculture	1	.2409	3.6622	-.1402
Business	1	.0170	.2593	.0166
Education	1	1.6095	24.5208**	.1856
Engineering	1	.3743	5.7026*	-.0900
Family Life	1	.0799	1.2172	.0655
HASS	1	.8510	12.9753**	.1441
Natural Resources	1	.1163	1.7711	-.0930
Science				-.0886
<u>Level</u>	6	.0837	1.2753	
100 Level	1	.0002	.0043	-.0022
200 Level	1	.0016	.0245	.0063
300 Level	1	.2407	3.6668	-.0669
400 Level	1	.2232	3.4000	-.0791
500 Level	1	.0474	.7215	-.0344
600 Level	1	.0070	.1062	-.0165
700 Level				.1928
Error	299	.0656	RSQ = .2056	

\*Significant at 5% level.

\*\*Significant at 1% level.

Whether a course is required or elective has negligible relation to an instructor's ratings.

The colleges of the University receive significantly different average ratings. Instructors in the colleges of Education, Family Life and Humanities, Arts and Social Sciences generally rate higher than other colleges. The colleges of Engineering and Science generally receive lower average ratings.

The level of a course an instructor teaches has little relation to instructor ratings on the USU FEQ.

### Discussion

Of concern to this author has been the careless and misinformed interpretation of faculty evaluation scores in some departments of the University. This negligence has come to her attention through discussion with instructors and administrators while administering the USU FEQ during the 1974-75 school year.

The results of this study raise some questions concerning the limitations of this instrument. These limitations must be considered when interpreting FEQ scores. The limitation most easily interpreted is that of college. Since mean scores differ significantly from college to college, scores should be analyzed within the context of the college under whose auspices a course is taught.

More difficult questions arise when examining the relationship between total mean score and question 23. It can be seen that question 23 would issue comparable results to the total form. How is such a relationship accounted for? One possibility is that total mean score reflects the integration of a number of variables that together compose good instruction and, the relationship reflects the same kind of process in students' minds when answering question 23. The students, then, are considering the specifics of quality teaching when making their assessment of instructors. If this is the case, one might feel comfortable about this relationship.

However, if the students are considering only the general likeableness of the instructor when making the comparison in question 23,

other questions arise. Does the relationship between question 23 and total mean score indicate that the FEQ assesses personality interactions between students and instructors rather than good instruction? Or, does the relationship between total score and question 23 reflect a similar relationship between the specifics of quality teaching and general likeableness of instructors? In other words, is a good teacher also a likeable one?

Consideration of such questions must be made in order to use the USU FEQ in the most productive manner. The effectiveness of the faculty evaluation program rests on the wisdom with which scores are interpreted.

#### Limitations

The outcomes of this investigation may be limited by any one or a combination of the following:

1. The results of this study were obtained from fall quarter evaluations. Ratings may relate to the quarter in which the course was evaluated.
2. The USU population was the only university used in the sampling. Results, therefore, will be generalizable to USU alone.
3. This study has been concerned with environmental variables, i.e., class size, level, college and required vs. elective classes. Student, class, or teacher characteristics have not been considered. Significant relationships may exist in these areas.
4. Within university policy, professors are allowed to choose the quarter in which they will be evaluated. Therefore, some volunteer

effects may be present in this study. However, extra encouragement to evaluate fall quarter was given to teachers to help control for these effects.

### Conclusions

From this study of the relationship between selected variables and total mean scores, the following may be concluded:

1. Question 23 could be used to obtain a quick, easy estimate of student ratings of an instructor.
2. The significant differences between ratings instructors in different colleges receive present an area of concern. When comparing instructors from different colleges, precautions may need to be taken. Differential norms for the colleges could be considered.
3. The size of an instructor's class is of no significant concern when interpreting FEQ results. The results of this study show no basis for the employment of differential norms for differing class sizes.
4. No basis for considering whether a course is required or elective and level of a course when interpreting questionnaire results is given.

### Recommendations

From the above limitations and conclusions the following recommendations are made:

1. The college under whose auspices an instructor teaches should be considered when interpreting FEQ results. College norm groups should be considered for inclusion on results printouts.



2. The significantly higher ratings received by instructors in the College of Education should be investigated. The following questions arise:

- A. Do students rate instructors higher if instructors give them higher grades? (Grades received in Education are higher.)
- B. Are the students in the college more empathetic to their instructors and, therefore, more lenient in their ratings?
- C. Is the course content less demanding or more interesting so as to receive higher student ratings?
- D. Are, in fact, the instructors in the College of Education better teachers? If so, what are the characteristics that make them better?

These questions and others should be researched.

3. Since relationships may vary according to the particular aspect of teaching performance that the student is asked to rate (Clark & Keller, 1954), a factor analytic study including the selected variables of this study and all of the questions on the USU FEQ may be profitable.

4. Scores on the USU FEQ may be related to the department in which the course is taught. Correlations might be determined for the departments which have a large enough N.

5. Determination of whether students' subjective criteria in rating faculty match the faculty members' goals in teaching may disclose some valuable information about the USU FEQ.

6. The benefits from student evaluation of instruction can only exist to the extent that ratings represent valid appraisals of classroom

instruction. Research must be conducted in answer to the question,

"Is the USU Faculty Evaluation Questionnaire a valid instrument?"

Literature Cited

- Bejar, I. A survey of selected administrative practices supporting student evaluation of instruction programs. Minneapolis: University of Minnesota, Measurement Services Center, 1973.
- Clark, K. E., & Keller, R. J. Student ratings of college teaching. In R. A. Eckert (Ed.), A university looks at its program. Minneapolis: University of Minnesota Press, 1954.
- Coffman, W. E. Determining student's concepts of effective teaching from their ratings of instructors. Journal of Educational Psychology, 1954, 45, 277-286.
- Cohen, J., & Humphreys, L. G. Memorandum to faculty. University of Illinois: Department of Psychology, 1960.
- Costin, F., Greenough, W. T., & Menges, R. J. Student ratings of college teaching: Reliability, validity and usefulness. Review of Educational Research, December 1971, 41(5), 511-535.
- Crawford, P. L., & Bradshaw, H. L. Perception of characteristics of effective university teachers: A scaling analysis. Educational and Psychological Measurement, 1968, 28, 1079-1085.
- Downie, N. W. Student evaluation of faculty. Journal of Higher Education, 1953, 23, pp. 495-496; 503.
- Eble, K. E. The recognition and evaluation of teaching. Salt Lake City, Utah: American Association of University Professors and Association of American Colleges, Project to Improve College Teaching, 1970.
- French, G. M. College students' concept of effective teaching determined by an analysis of teacher ratings. Dissertation Abstracts, 1957, 17, 1380-1381.
- Gadzella, B. M. College student views and ratings of an ideal professor. College and University, 1968, 44, 89-96.
- Gage, N. L. The appraisal of college teaching. Journal of Higher Education, 1961, 32, 17-22.
- Goodhartz, A. S. Student attitudes and opinions relating teaching at Brooklyn College. School and Society, 1948, 68, 345-349.
- Gustad, J. W. Policies and practices in faculty evaluation. Educational Record, 1961, 42, 194-211.

- Guthrie, E. R. The evaluation of teaching: A progress report. Seattle: University of Washington, 1954.
- Heilman, J. D., & Armentrout, W. D. The rating of college teachers on ten traits by their students. Journal of Educational Psychology, 1936, 27, 197-216.
- Kent, L. Student evaluation of teaching. In American Council on Education, Handbook of selected readings in evaluation of teaching performance. Office of Instructional Development, Colorado State University, 1967, pp. 56-85.
- Kohlman, R. G. A comparison of faculty evaluations early and late in course. Journal of Higher Education, 1973, 44(8), 587-595.
- Lovell, G. D., & Haner, C. F. Forced-choice applied to college faculty rating. Educational and Psychological Measurement, 1955, 15, 291-304.
- McDaniel, E. D., & Feldhusen, J. F. Relationships between faculty ratings and indexes of service and scholarship. Proceedings of the 78th Annual Convention of the American Psychological Association, 1970, 5, 619-620.
- Muscella, D., & Rusch, R. Student opinion on college teaching. Improving College and University Teaching, 1968, 16, 137-140.
- Remmers, R. H., & Elliot, D. N. The Indiana College and University Staff-Evaluation Program. School and Society, 1949, 70, 168-171.
- Smith, A. A. College teachers evaluated by students. Sociology and Social Research, 1944, 28, 471-478.
- Solomon, D. Teacher behavior dimensions, course characteristics, and student evaluations of teachers. American Educational Research Journal, 1966, 3, 35-47.
- Spencer, R. E. The Illinois Course Evaluation Questionnaire Manual of Interpretation. University of Illinois, Measurement and Research Division, Office of Instructional Resources, 1968.
- Stewart, C. T., & Malpass, L. F. Estimates of achievement and ratings of instructors. Journal of Educational Research, 1966, 59, 347-350.
- Treffinger, D. J., & Feldhusen, J. F. Predicting students' ratings of instruction. Proceedings of the 78th Annual Convention of the American Psychological Association, 1970, 5, 621-622.

U. S. President's Commission on Campus Unrest. The report of the President's commission on campus unrest. New York: Arno Press, 1970.

Zelby, L. W. Student-faculty evaluation. Science, March 1974, 183, 1267-1270.

Appendices

Appendix A

USU Faculty Evaluation Form





# Utah State University Course Evaluation

This course is 43  
 A. required in my major.  
 B. required in my minor.  
 C. part of my general education.  
 D. a free elective. A B C D

## INSTRUCTIONS

In the spaces provided (at the left) write the name of the instructor being evaluated and the course number. Use a soft lead pencil to mark your answers in the appropriate space.

Use the following alternatives for questions 1 through 16.

NA – Statement is not applicable to this course.

SA – Strongly agree with the statement.

A – Agree with the statement.

D – Disagree with the statement.

SD – Strongly disagree with the statement.

- |  |             |
|--|-------------|
| 1. The purpose and goals of this course have been clearly stated.  | NASA A D SD |
| 2. The stated course objectives correspond closely to what has actually been done.   | NASA A D SD |
| 3. Class time was well used.   | NASA A D SD |
| 4. The in-class and out-of-class activities complemented each other without undue repetition.  | NASA A D SD |
| 5. The instructor was genuinely concerned with the students who were having difficulty.  | NASA A D SD |
| 6. Comments on written exams and assignments were fair and helpful.  | NASA A D SD |
| 7. Students were allowed and encouraged to ask questions and to express their opinions.  | NASA A D SD |
| 8. Early in the course the students were informed of the way their performance would be evaluated and they were evaluated accordingly. | NASA A D SD |
| 9. The instructor summarized well and made major points easy to identify.  | NASA A D SD |
| 10. The instructors presentations were well organized and easy to follow.  | NASA A D SD |
| 11. The instructor has the ability to stimulate and to develop a "thirst for knowledge".   | NASA A D SD |
| 12. The instructor was enthusiastic about the course.  | NASA A D SD |
| 13. The instructor had control of the discussions and his supervision of the class was good.   | NASA A D SD |
| 14. An effective learning environment was maintained.  | NASA A D SD |
| 15. I am more interested in the subject matter now than I was at the beginning of the course.  | NASA A D SD |
| 16. I was interested in the course when I registered for it.   | NASA A D SD |

Use the following alternatives for questions 17 through 22.

NA – Not applicable.

E – Excellent.

G – Good

F – Fair.

P – Poor.

Please rate each of the following.

- |   |            |
|---|------------|
| 17. Textbooks.  | NA E G F P |
| 18. Supplementary reading.  | NA E G F P |
| 19. Examinations.   | NA E G F P |
| 20. Laboratories.   | NA E G F P |
| 21. Lectures.   | NA E G F P |
| 22. Class discussions.  | NA E G F P |
| 23. Compared to the other instructors I have had, this instructor is in the |            |
| A. Top 10%  |            |
| B. Top 25%  |            |
| C. Middle   |            |
| D. Bottom 25%   |            |
| E. Bottom 10%   | A B C D E  |

Now please respond to the questions on the back of this sheet.

COURSE NUMBER:

Appendix B

Memo to USU Faculty

## Utah State University

## Faculty Evaluation 1975-'76

Faculty evaluation forms are now available in the Counseling and Testing Center, room 2, Old Main. It is the instructors responsibility to request forms for each of the classes to be evaluated. The forms will be mailed to the instructor or can be picked up in room 2. The instructor starts the administration of the form and then turns it over to a class member or secretary to return them to Counseling and Testing. The forms are then processed by the computer and printouts of the results are made. Results from faculty evaluations are returned to the instructor after the grades for the class being evaluated are turned in. A computer printout is returned to the faculty member, the department head, the college dean, and the student files.

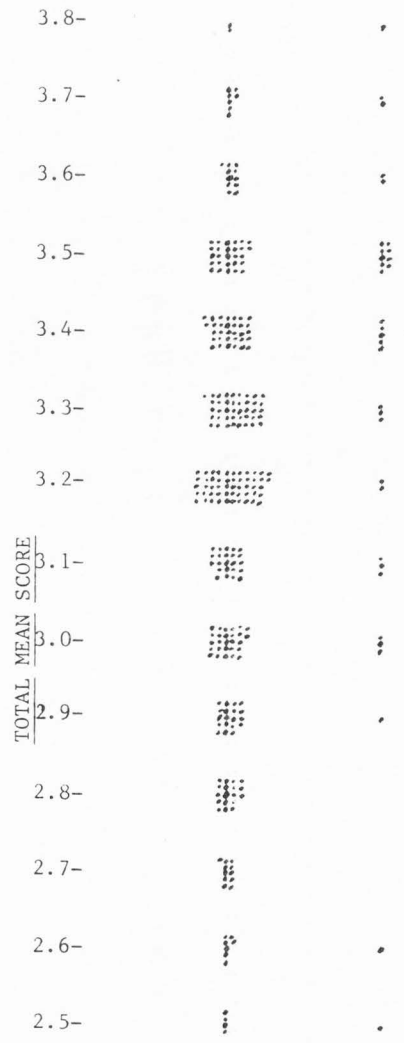
USU's current policy concerning faculty/student evaluation is as follows. Over the period of three years, a faculty member will be evaluated in every undergraduate course taught that has a class size of ten or more. Each faculty member teaching undergraduate classes must be evaluated in at least one course every year. The faculty member may choose the course to be evaluated. Faculty members may have as many courses as they wish evaluated. Course evaluation will not be administered before the sixth week of the quarter and should not be administered after the ninth week in the quarter.

The norms for the results have been updated and expanded. When the original base for the deciles was established, it was very limited. The base is now larger, more diversified and should give more accurate comparisons.

Appendix C

Scatter Diagrams Comparing Total Mean Scores  
and the Selected Variables

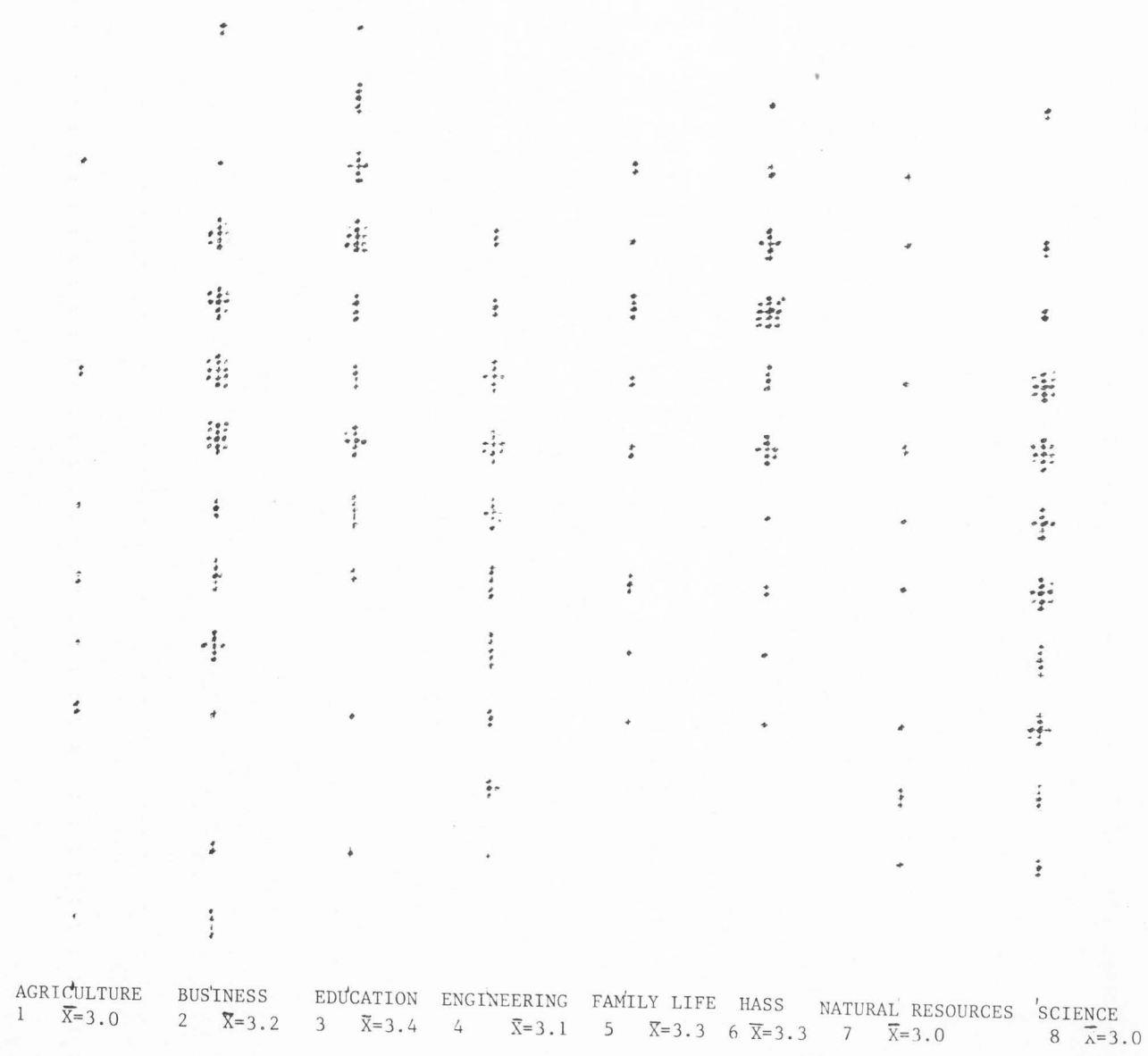
TOTAL MEAN SCORE	% COMPARISON									
	1.2-	1.3-	1.4-	1.5-	1.6-	1.7-	1.8-	1.9-	2.0-	2.1-
2.5-	.	.	.	.	.	.	.	.	.	.
2.6-	.	.	.	.	.	.	.	.	.	.
2.7-	.	.	.	.	.	.	.	.	.	.
2.8-	.	.	.	.	.	.	.	.	.	.
2.9-	.	.	.	.	.	.	.	.	.	.
3.0-	.	.	.	.	.	.	.	.	.	.
3.1-	.	.	.	.	.	.	.	.	.	.
3.2-	.	.	.	.	.	.	.	.	.	.
3.3-	.	.	.	.	.	.	.	.	.	.
3.4-	.	.	.	.	.	.	.	.	.	.
3.5-	.	.	.	.	.	.	.	.	.	.
3.6-	.	.	.	.	.	.	.	.	.	.
3.7-	.	.	.	.	.	.	.	.	.	.
3.8-	.	.	.	.	.	.	.	.	.	.
3.9-	.	.	.	.	.	.	.	.	.	.
4.0-	.	.	.	.	.	.	.	.	.	.



REQUIRED	ELECTIVE
1	0
$\bar{X}=3.2$	$\bar{X}=3.3$

3.8-  
3.7-  
3.6-  
3.5-  
3.4-  
3.3-  
3.2-  
3.1-  
3.0-  
2.9-  
2.8-  
2.7-  
2.6-  
2.5-

TOTAL MEAN SCORE



COLLEGES

C-3

CLASS LEVEL	100-199	200-299	300-399	400-499	500-599	600-699	700-799
3.8-	.						
3.7-	••	•	•	•	•	•	•
3.6-	••••	••••	••••	••••	••••	••••	••••
3.5-	•••••	•••••	•••••	•••••	•••••	•••••	•••••
3.4-	••••••	••••••	••••••	••••••	••••••	••••••	••••••
3.3-	•••••••	•••••••	•••••••	•••••••	•••••••	•••••••	•••••••
3.2-	••••••••	••••••••	••••••••	••••••••	••••••••	••••••••	••••••••
3.1-	•••••••••	•••••••••	•••••••••	•••••••••	•••••••••	•••••••••	•••••••••
3.0-	••••••••••	••••••••••	••••••••••	••••••••••	••••••••••	••••••••••	••••••~
2.9-	•••••••••••	•••••••••••	•••••••••••	•••••••••••	•••••••••••	•••••••••••	•••••••••••
2.8-	••••••••••••	••••••••••••	••••••••••••	••••••••••••	••••••••••••	••••••~	••••••••••••
2.7-	•••••••••••••	•••••••••••••	•••••••••••••	••••••••••~	•••••••••••••	•••••••••••••	•••••••••••••
2.6-	••••••••••••••	••••••••••~	•••••••••••••	•••••••••••••	•••••••••••••	•••••••••••••	•••••••••••••
2.5-	•••••••••••••••	•••••••••••••	•••••••••••••	••••••••••~	•••••••••••••	•••••••••••••	•••••••••••••
TOTAL MEAN SCORE	$\bar{X}=3.2$	$\bar{X}=3.2$	$\bar{X}=3.1$	$\bar{X}=3.2$	$\bar{X}=3.2$	$\bar{X}=3.3$	$\bar{X}=3.5$



