This annotated bibliography on student evaluation of faculty contains articles published between 1965 and 1970. The 107 articles are divided into the following categories: surveys; factor analytic surveys (usually authors of these studies address themselves to different questions than do authors of non-factor analytic surveys); experiments (in these studies, 1 or more independent variables is actively manipulated); reviews; and non-empirical articles. (JS)
Evaluation of instruction is primary to the functioning of an educational institution. Although there are many methods to evaluate teaching, much controversy has revolved around the development of instruments to measure student evaluation of instruction.

Coverage. This report covers articles concerning college student evaluation of faculty members published between January, 1965 and December, 1970. Four sources were searched: Research in Education, Current Index to Journals in Education, Education Index, and Psychological Abstracts. The articles located through these sources were then examined for additional references.

Organization. The articles are categorized into five groups:
1) Surveys--Differential observation. Includes quasi experiments in which an independent variable is not actively manipulated.
2) Factor Analytic Surveys--same as 1) above, but factor analysis is used in the analysis of the data. In consequence, the authors of these studies usually address themselves to different questions than the authors of non-factor analytic surveys do.
3) Experiments--One or more independent variables is actively manipulated.
4) Reviews--Articles which include reference to published data by other investigators.
5) Non-empirical articles--Only a sampling of these unimportant articles is included. None are annotated.

Referencing. Articles which were not published in a journal or book include a six-digit ED number in the reference. These articles can be obtained (in microfiche or hard cover) by writing:

ERIC Document Reproduction Service
Leasco Information Products
4827 Rugby Avenue
Bethesda, Maryland 20014

Doctoral dissertations. Ph.D. theses are included but unannotated because they were not examined.

1 Thanks are sincerely expressed to Arlene Tervakoski for typing this paper.
2 Deborah Lorenson's assistance in the library search is gratefully acknowledged.
SURVEYS


In the spring of 1966, the American Council on Education initiated a survey to ascertain current techniques for the evaluation of undergraduate instruction. The authors report a detailed analysis of the major results of this survey. A questionnaire was sent to the deans of the entire population of higher educational institutions. Among the 1,110 usable questionnaires, the most frequently indicated source of information for evaluation of teaching effectiveness was evaluations by the dean and department chairmen. Systematic student ratings were poorly developed.


Reports on a booklet resulting from student opinion of courses and faculty replies. Except for a few percentages, no data. Incomplete sampling information.


Found no significant differences in student ratings between student types (assessed by Myers-Briggs Type Indicator), but student opinion differed on different faculty types.


Governmental support and publication frequency was found to be directly related to rankings of teacher effectiveness.


Reports three studies showing how similar several groups of students rate teacher traits.

Carney, R.E. and McKeachie, W.J. Personality, sex, subject matter and student ratings. *Psychol. Record*, 1966, 16 (2), 137-144.

A good study with replications across courses, institutions, and different instructors and student backgrounds.

Students rated verbal descriptions of three "types" of instructors.


Describes the relationships between a seminar in the teaching of psychology and later student ratings of the participants.


Compares ratings of teachers by faculty members, university administrators, and university students.


Questionnaire given to author's students. Incomplete description of procedures. No data.


Uncommonly poor study.


Poorly reported study.


Replication and extension of Yamamoto and Dizney (1966).


Careful development of a student generated evaluation form.

Description of a well-established program of student ratings, including some data.


An incompletely reported study which purports to show that high authoritarian students are more critical of instruction.


Most relationships found between ratings and student related characteristics were small.


Q sorts used to assess perceptions.


Examines the relationships between students' gross subjective judgments and their responses to a typical rating scale.


Published ratings produced as a result of the survey.

Hussain, K.M. and Leestamper, R. Survey on criteria of teaching effectiveness at New Mexico State University, June 1968, 53p. ED 023 365.

The survey spans four universities as well as faculty, students, and alumni. All groups agreed that the most important criterion was "being well prepared for class".


A poorly described study which compares the evaluation of possible professor roles (research, teaching, and character development) by college freshmen and their parents.

This article reports on the important survey initiated by the American Council on Education as described by Astin and Lee (1966).


Compares the relative importance of 12 areas in instruction among faculty and students at a community college.

Lathrop, R.G. Unit factorial ratings by college students of courses and instructors, 1968, 12p. ED 028 686.

Found that perceived learning is one of the factors significantly influencing ratings of faculty. One of the few articles in the literature to offer a theoretical framework.


Found a generally favorable image. Demographic characteristics of the respondents did not affect the ratings.

Musella, D. and Rusch, R. Student opinion on college teaching. *Imp. Coll. Univ. Teach.*, 1968, 16 (6), 137-140.

Mailed a questionnaire to college seniors in order to identify professors considered to have improved student thinking and to obtain opinions of behaviors that promote thinking.

Orange Coast College. Instructor rating scale study, March 1969, 6p. ED 028 775.

Brief report of the mean ratings of 26 instructors.


Found negligible relationships between ratings of instructors and student, instructor and situational (e.g. hour of class) characteristics.


Poorly reported study drawing unwarranted conclusions.

Two-page report of the ranking of 60 criterion behaviors developed from categorizing 13,643 behaviors.


One of several studies by Pervin which uses TAPE (Transactional Analysis of Personality and Environment). Develops an interactive model to explain students' perceptions of themselves, the faculty, and the college.


Similar to Pervin's other studies.


Reports Negro-white differences in what the respondents expect of teachers.


Respondents ranked ten descriptive statements.


Found that student evaluations of a reading course tend to be more closely related to evaluations of the teacher than to any measured improvement in reading.


Concludes that student ratings are relatively unbiased since they are not much related to either student or instructor characteristics.


Found that specific need patterns (on Edwards Personal Preference Schedule) are not associated with specific patterns of perception of instructor (on Purdue Rating Scale for Instruction). However, a sex x need interaction was found.

Presents a rating form, instructions, and the ratings received by specific faculty members.


Describes a student rating program involving recorded class sessions, conferences, and class discussions. The SET program is viewed as producing raw data for application of the critical incident technique.


On the basis of GPA, a High Achiever and a Low Achiever group were developed. The two groups differed in responses to instruction evaluation forms.


A course evaluation questionnaire, based upon 50 items and used at 13 different institutions, is briefly described.

Stallings, W.A. and Singhal, S. Some observations on the relationships between research productivity and student evaluations of courses and teaching, February 1969, 10p. ED 027 838.

Report significant, but low correlations which provide some support for the assertion that a productive researcher tends to be a good teacher.


Contrary to the bulk of the research, the authors reported that students expecting high grades rated their instructors significantly higher than did those expecting low grades. Year in school was also found to be correlated with the ratings.


Respondents were given the same Likert rating scale under two different sets. Seven differences were found, e.g. agreed that students take advantage of permissive professors (real) but this is not an appropriate thing to do (ideal).

Report adjective sorts for real and ideal teachers.


Report on the development and use of an evaluation questionnaire. A larger percentage of graduate students reacted favorably to most items than undergraduates.


Student preferences for different types of college professors were studied. Teacher, researcher, socialite, and administrator were preferred in that order. In agreement with most of research, no sex differences in responding were obtained.


**Doctoral Dissertations**


Sorey, K.E. A study of the distinguishing personality characteristics of college faculty who are superior in regard to the teaching function. *Dissert. Abs.*, 1968, 28 (12-A), 4916.

Walker, B.D. An investigation of selected variables relative to the manner in which a population of junior college students evaluate their teachers. *Dissert. Abs.*, 1969, 29 (9-B), 3474.

Warren, P.B. A study of lower class and middle-upper class students' perception of the behavior traits of the effective teacher. *Dissert Abs.*, 1968, 29 (2-A), 523.

FACTOR ANALYTIC SURVEYS


A factor analysis of the responses of 139 students resulted in five stable factors: teaching ability, feedback to students, negative attitudes, student overload, and structure. Personal qualities of the teacher, sex of the student, grades in the class, and overall GPA did not influence student assessment of teacher performance.


Applies Pervin's concept of person-environment fit to learning in college classes. The aspect of the person studied was the student's perception of an Ideal Teacher Style. The aspect of the environment considered was the student's perception of his instructor's Teacher style. The congruence between the ratings of these two instruments did not predict learning as well as the ratings of Teacher Style alone.


The central underlying hypothesis of several of Kerlinger's studies is that perceptions of the traits of effective teachers are, in part, a function of the student raters' attitudes toward education. A factor analysis of a Q sort instrument resulted in three factors: progressive, traditional, and an unnamed factor.


A carefully executed study which replicates factor structures across several samples from different schools. The three named first order factors included positive person orientation, systematic task organization, and functional flexibility. The second order factors were not successfully replicated.


Found that those students who characterized themselves as having progressive philosophies of education tended to perceive person-oriented traits as desirable for teachers, whereas traditionists selected task-oriented traits as desirable. Two second order factors were found: progressivism and traditionalism.


The responses of 1,097 students to two instructor evaluation forms were factor analyzed. Two of the nine factors (instructor impact and instructional impact) accounted for 64 per cent of the rotated variance.


Applied the TAPE analysis to 3,016 students from 21 colleges. It was found that the discrepancies between student perceptions of themselves and their college were related to dissatisfaction with college. Information concerning the reliability, validity, and factorial structure of TAPE is included.


Eleven types of teachers were rated by 65 students on 20 semantic differential scales. A factor analysis yielded seven labeled factors: socio-economic, racial, social studies aptitude, class in school, masculine sophistication, social disposition, and emotional instability. The conceptions students have of teachers are influenced, in part, by the personality traits and background experiences of the students themselves.


The responses of the students in the classes of 229 teachers (total N is not reported) to a student questionnaire were factor analyzed. The 10 factors extracted were similar to those found in an earlier study by the author.


A well executed study based upon Fiedler's cubic leadership model which postulates an interaction between the characteristics of the leader, the characteristics of the group, and the complexity of the tasks. In agreement with the model, the authors found that different types of students prefer different types of instructors.

An interesting study in which 227 students taught by three instructors completed an instruction rating form and a personality test. Correlations between factor scores and the personality test indicated that in some cases, the same personality variable correlated with a given factor score in opposite directions from one instructor to another. The authors develop an interactive point of view.

Doctoral Dissertations


EXPERIMENTS


A rating scale was given to two classes taught by the author. One class was Psychological Statistics which the author had taught before and found enjoyable and the other class was Introductory Psychology which the author felt unprepared for and disinterested. Predicted differences were obtained.


Students were administered evaluation forms under one of four instructional conditions. The instructional conditions had no effect on the forced-choice ratings but had a significant effect on the graphic ratings. It was concluded that the forced-choice scale is resistant to bias in student ratings.


A more complete description of the well executed study by Sharon (1970).


A study employing appropriate control groups in which it was found that students preferred patterns of teacher influence encouraging participation regardless of ability, sociometric status, and personality traits.

Doctoral dissertation.

REVIEWS


Three approaches to instructor evaluation are discussed: student evaluation by questionnaire, instructor self-appraisal, and team evaluation.


A monograph which includes 1) a discussion and appraisal of methods currently used, and 2) a case for changing the purposes, methods and criteria of faculty assessment. 128 ref.


A short discussion citing seven references.


Teacher evaluation is discussed within a broad context of adult learning. Includes 24 references, Allport's teacher evaluation plan, a self-evaluation check-list and an appendix of teaching techniques.

Eble, K.E. The recognition and evaluation of teaching. Project to Improve College Teaching, Salt Lake City, Utah, 1970.

Includes the how's and why's of improvement of teaching, student evaluation instruments, and the relationship between student evaluation and faculty review. Six unpublished studies are included in the appendix and an extensive, two-part bibliography is attached.

Eidell, T.L. and Kleve, J.A. Annotated bibliography on the evaluation of educational programs, November 1968, 19p. ED 025 857.

Only a few references to student evaluation of instruction are included.


Presents a selected sample (11) of empirical studies pertaining to the perceived desirable characteristics of the effective college teacher.

A paper developed as a requirement for a college course. Includes a series of excerpts concerning policies and procedures for instructional evaluation at 23 California junior colleges.


Includes about 100 unannotated references (1960-1966) from several areas, one of which is teacher evaluation.


Chapters 2 and 4 include a simplified discussion of teacher evaluation. Old references cited.


Includes a nine-page discussion, three appendices and 30 selected references.


A cookbook essay on how to evaluate, why and then what to do. Includes recommended guidelines for teacher evaluation and a 40-item student opinion questionnaire as appendices. 41 pre-1965 ref.


Includes 56 entries (1961-1968) dealing with methods, procedures, and problems involved in faculty evaluation by students, administrators, and peers.


A very useful article in which the author examines the same misconceptions and studies as he did in a 1957 article. A 17-item form is attached.


Chapter 22 includes a discussion based upon the author's 1957 article. Extensive pre-1965 bibliography.

Discuss four techniques to study the criterion problem, misconceptions regarding evaluation, and microteaching (teaching for short periods and immediate feedback).


The author lists 18 conclusions presumably based upon his research and a literature review, but the conclusions are not documented separately.


Includes summaries of successful and proposed criteria for measuring faculty performance and barriers to their use. 22 ref.


Deals with objections to student evaluation of faculty as McKeachie (1969) did. Also includes a comparison of two factor analytic studies: when similar items are chosen, similar factors are revealed.


Authors' basic tenet is that a definition of teaching success which is formulated in terms of some single fixed teacher ideal is both untenable and inappropriate, i.e. different raters have different expectancies. Some good ideas. 29 ref.
NON-EMPIRICAL ARTICLES


Bryant, P.T. By their fruits ye shall know them. J. Higher Educ., 1967, 38 (6), 326-330.


