

THE SUPERIOR COLLEGE TEACHER FROM THE STUDENTS' VIEW

Kenneth A. Feldman, *Department of Sociology, State University of New York at Stony Brook*

.....

This analysis attempts a systematic synthesis of the rather large body of research on college students' views on teaching and their assessment of the effectiveness of various attitudes, behaviors, and pedagogical practices of instructors. Across studies, the following characteristics are consistently associated with superior college teachers or teaching (as determined in a variety of ways): stimulation of interest; clarity and understandableness; knowledge of subject matter; preparation for, and organization of, the course; and enthusiasm for the subject matter and for teaching. Friendliness (concern and respect for students), helpfulness (availability), and openness to others' opinions (encouragement of class questions and discussion) are characteristics that students say they prefer in teachers (especially when they freely describe their ideal or best teacher and the characteristics they see as important to good teaching). However, a teacher's standing on these characteristics appears not to be particularly important to students' global assessment of their actual teachers on teacher-evaluation questionnaires. Finally, relative to other characteristics, certain regulative activities of the instructor are typically and consistently lower in importance for superior teaching, at least by the several indicators used in the studies under review. The analysis concludes with interpretations and cautions.

.....

Key words: effective college teaching; descriptions of the ideal teacher; evaluation of college teachers; course evaluation; student ratings

Analysts and practitioners in the field of higher education have recurrently sought to specify those attitudinal and behavioral characteristics of college instructors that constitute effective teaching and that distinguish the superior teacher. This goal is not an easy one, for the conflicting issues, values, theoretical orientations, and methodologies surrounding the topic are complex, and the sources of relevant information

are many (see, *inter alia*, Axelrod, 1973; Brandis, 1964; Centra, 1972; Doyle, 1975; Eble, 1972; Mann and associates, 1970; Miller, 1972, 1974; and Neeley, 1968). One source of data, of course, is college students themselves—those persons having the most “direct access to the professor’s wares,” as Riley et al. (1950) have put it. Students, incidentally, have been overly praised perhaps as often as they have been ignored or even maligned as a source of opinion about teachers and teaching. Polemics aside, there exists considerable research on college students’ views on teaching and their assessment of the effectiveness of various attitudes, behaviors, and pedagogical practices of their instructors. The present analysis attempts a rather systematic synthesis of this research, reviewing those studies that deal exclusively or primarily with undergraduate students at North American and Canadian colleges and universities.

CHARACTERISTICS REPORTED BY STUDENTS AS ASSOCIATED WITH IDEAL (BEST) TEACHERS AND AS IMPORTANT TO EFFECTIVE TEACHING

There are different ways in which college students have been asked to specify the attitudes and behaviors they feel are most important for superior teaching and most characteristic of the superior teacher. In some studies, students have been asked to describe their ideal teacher. In other studies, the request to students has been to indicate the characteristics that they feel are especially important to good teaching. A third, somewhat less direct approach has been to have students describe the best teachers they have had.

The array of characteristics that have appeared within and across these studies can be classified into a smaller number of categories or dimensions. The dimensions used in the present analysis, along with examples of responses falling into these categories, are given in Table I.¹ For the studies under consideration in this section of the analysis (see Appendix A), each characteristic in a study could usually be coded into one of these dimensions. Sometimes, however, it was necessary to code a characteristic as belonging to more than one category (see Appendix A). For example, of the “criteria describing an ideal professor” listed in a study by Gadzella (1968a), the criterion identified as “Encourages Students’ Participation (provides for questioning periods, encourages independent thinking)” is coded here as categories no. 16 and no. 17.

Each of the studies on which Table I is based has information about differences in the importance to students of various attitudes, behaviors, and teaching practices of instructors: Either these characteristics are presented in the order of their importance in the study or data from which a

rank ordering can be derived are given (see Appendix A). Since the number of characteristics that are ranked varies from one study to another, direct comparison of the ranks of a given characteristic in different studies would be problematic. To rank 8 out of 25 characteristics is clearly not equivalent to ranking 8 out of 8 characteristics. Thus, in order to increase comparability among studies, the ranks of the characteristics were standardized in each study by dividing each rank by the number of characteristics in the study.² The smaller this fraction, the greater the rank-ordered importance of the characteristic. When more than one of the characteristics in a given study was coded into the same category, the standardized ranks for these characteristics were averaged to produce the standardized rank for that coding category for that study. For example, of the 25 behavioral characteristics that students at the University of Toledo felt most important to effective teaching, being well prepared for class ranked in first place and organizing the course in logical fashion ranked eighth (see Perry, 1969). Since both of these are coded as part of the same dimension in the present analysis (category no. 5), the standardized rank for this particular dimension for this one study is the average of .04 (i.e., $1/25$) and .32 (i.e., $8/25$) or .18 (i.e., $[\.04 + .32] / 2$).

The studies and their results have been divided by whether students were directed to specify the characteristics of the ideal teacher, the characteristics important to good teaching, or the characteristics of the students' best teachers. (The exact directive to the student varies, of course, from study to study, as may be noted in the short descriptions of the studies given in Appendix A.) The studies were further divided—in what turned out to be the more significant division—into those studies in which students could freely respond with a list of characteristics of their own choosing (hereafter referred to as the nonstructured-response studies) and those studies in which students were given a preset list of characteristics for their consideration (the structured-response studies). Within each of the resultant six divisions, and for each of the dimensions under study, an overall average standardized rank (ASR) was calculated (given in Table I). These overall average standardized ranks, in turn, were themselves ranked from high to low. The ten most important dimensions, as indicated by the ranks of the overall average standardized ranks, have been roughly ordered and clustered for the nonstructured-response set of studies and for the structured-response set of studies (see the first two sections of the three sections into which Table III is partitioned). The three subdivisions within each of these two sets of studies (cols. 1-3 and cols. 4-6) are considered together, since the differences in overall average standardized rank of the characteristics are smaller when the three subsets of studies are compared (within the two larger sets of studies) than when

the results of the nonstructured-response sets and the structured-response sets as a whole are compared.

TABLE I. Characteristics of Ideal and Best College Teachers, and Characteristics Important to Superior College Teaching, as Seen by College Students

Note: In this table, nonstructured response refers to studies in which students were free to list characteristics of their own choosing; structured response refers to studies in which students were asked to rate or rank a predetermined list of characteristics.

Ideal refers to studies in which students listed or rated characteristics of the ideal teacher. Important refers to studies in which students listed or rated characteristics they felt to be important to good teaching. Best refers to studies in which students listed or rated characteristics of their best teachers.

ASR is an abbreviation for overall average standardized rank, i.e., the average of the standardized ranks for a particular dimension (or category of characteristics) across studies (see text). If only one study was found in which a particular dimension appeared, the standardized rank (ASR) is the range of standardized ranks and a listing of the studies on which the ASR is based. These studies are ordered from high to low with respect to their standardized ranks for the dimensions under consideration.

S1 refers to study no. 1, S2 refers to study no. 2, and so on through S49 (study no. 49). The studies are identified and briefly described in Appendix A.

1. Stimulation of interest: "the instructor puts material across in an interesting way"; "the instructor gets students interested in the subject"; it was easy to remain attentive"; "the teacher stimulated intellectual curiosity"; etc.

Nonstructured response

Ideal: ASR = .28 (Range: .28-.29) S49, S42
 Important: ASR = .36 (Range: .09-.67) S17, S8, S2
 Best: ASR = .33 (Range: .08-.60) S6, S21, S10, S1, S22, S20, S45

Structured response

Ideal: ASR = .14 (Range: .10-.30) S30, S12, S13, S14, S31, S32, S43
 Important: ASR = .28 (Range: .10-.52) S4, S44, S3, S16, S27, S11, S15, S28,
 S29, S18, S26
 Best: ASR = .26 (Range: .19-.40) S47, S5, S19, S25

2. Instructor's enthusiasm (for subject or for teaching): "the instructor shows interest and enthusiasm in the subject"; the instructor seems to enjoy teaching"; etc.

Nonstructured response

Ideal: (.22) S42
 Important: ASR = .53 (Range: .07-.94) S33, S7, S49, S8, S2
 Best: ASR = .50 (Range: .21-.75) S6, S48, S20, S1, S45, S21, S35, S10, S22

Structured response

Ideal: ASR = .18 (Range: .06-.60) S14, S12, S13, S31, S32, S39, S38
 Important: ASR = .35 (Range: .08-.76) S37, S23, S11, S9, S36, S44, S24, S4, S34,
 S40, S3, S26
 Best: ASR = .42 (Range: .19-.93) S47, S46, S5, S19

3. Instructor's knowledge of subject matter: "the instructor has a good command of the subject material"; "the teacher has a thorough knowledge, basic and current, of the subject"; "the instructor has good knowledge about or beyond the text book"; etc.

Nonstructured response

Ideal: ASR = .23 (Range: .12-.33) S42, S49
 Important: ASR = .24 (Range: .04-.33) S8, S41, S17, S7, S2, S33
 Best: ASR = .33 (Range: .04-.67) S21, S10, S20, S45, S48, S6, S22, S1

Structured response

Ideal: ASR = .18 (Range: .04-.70) S12, S13, S31, S32, S38, S43, S14, S30, S39
 Important: ASR = .26 (Range: .07-.88) S16, S9, S27, S29, S34, S3, S28, S37,
 S26, S36, S40, S4, S15, S23, S18, S24
 Best: ASR = .11 (Range: .07-.16) S47, S55, S46

4. Instructor's intellectual expansiveness (and intelligence): "the instructor is well informed in all related fields"; "the instructor has respect for other subject areas and indicates their relationship to his or her own subject of presentation"; "the instructor exhibited a high degree of cultural attainment"; etc.

Nonstructured response

Ideal: ASR = .74 (Range: .65-.83) S42, S49
 Important: ASR = .77 (Range: .36-.96) S8, S17, S33, S2, S7
 Best: ASR = .69 (Range: .40-.86) S45, S6, S48, S21, S1, S20

Structured response

Ideal: ASR = .57 (Range: .53-.62) S14, S13, S32, S12, S31
 Important: ASR = .35 (Range: .20-.58) S26, S9, S3
 Best: (1.00) S25

5. Instructor's preparation; organization of the course: "the instructor was well prepared for each day's lecture"; "the presentation of the material is well organized"; "the overall development of the course had good continuity"; etc.

Nonstructured response

Ideal: ASR = .56 (Range: .32-.78) S42, S49
 Important: ASR = .63 (Range: .40-1.00) S33, S2, S7, S8
 Best: ASR = .44 (Range: .04-.88) S21, S22, S45, S35, S1, S20, S6, S48

Structured response

Ideal: ASR = .29 (Range: .16-.40) S12, S39, S31, S14, S43, S32, S13, S38
 Important: ASR = .29 (Range: .13-.90) S15, S18, S23, S9, S37, S34, S3, S27,
 S40, S24, S28, S36
 Best: ASR = .48 (Range: .26-.72) S45, S48, S19

6. Clarity and understandableness: "the instructor made clear explanations"; "the instructor interprets abstract ideas and theories clearly"; "the instructor makes good use of examples and illustrations to get across difficult points"; "the teacher effectively synthesizes and summarizes the material"; etc.

Nonstructured response

Ideal: ASR = .17 (Range: .11-.22) S49, S42
 Important: ASR = .60 (Range: .41-1.00) S7, S2, S33, S41
 Best: ASR = .39 (Range: .06-.82) S10, S1, S20, S22, S35, S48, S45, S6

Structured response

Ideal: ASR = .21 (Range: .10-.40) S30, S14, S31, S32, S12, S13, S38, S39
 Important: ASR = .22 (Range: .08-.50) S24, S11, S36, S40, S44, S15, S18, S37,
 S3, S23, S34, S26, S27, S28
 Best: ASR = .42 (Range: .25-.70) S19, S37, S25, S46

7. Instructor's elocutionary skills: "the instructor has a good vocal delivery"; "the teacher speaks distinctly, fluently and without hesitation"; "the teacher varied the speech and tone of his or her voice"; etc.

Nonstructured response

Ideal: (.74) S42
 Important: ASR = .45 (Range: .27-.88) S33, S7, S2, S17, S8, S41
 Best: ASR = .52 (Range: .40-.63) S20, S6

Structured response

Ideal: ASR = .74 (Range: .52-1.00) S31, S14, S13, S32, S12, S39, S38
 Important: ASR = .72 (Range: .44-1.00) S26, S9, S24, S34, S40, S3, S27, S23, S36
 Best: ASR = .63 (Range: .44-.73) S46, S19, S47

8. Instructor's sensitivity to, and concern with, class level and progress: "the instructor was skilled in observing student reactions"; "the teacher was aware when students failed to keep up in class"; "the instructor teaches near the class level"; "the teacher takes an active personal interest in the progress of the class and shows a desire for students to learn"; etc.

Nonstructured response

Ideal: —
 Important: (.74) S7
 Best: ASR = .65 (Range: .33-.89) S6, S20, S35

Structured response

Ideal: ASR = .19 (Range: .16-.24) S31, S32, S12, S13
 Important: ASR = .20 (Range: .08-.34) S24, S23, S37
 Best: ASR = .40 (Range: .16-.64) S46, S19

9. Clarity of course objectives and requirements: "the purposes and policies of the course were made clear to the student"; "the instructor gave a clear idea of the student requirements"; "the teacher clearly defined student responsibilities in the course"; etc.

Nonstructured response

Ideal: (.32) S42
 Important: (.60) S8
 Best: (.54) S35

Structured response

Ideal: —
 Important: —
 Best: ASR = .39 (Range: .25-.52) S19, S46

10. Nature and value of the course material (including its usefulness and relevance): "the instructor has the ability to apply material to real life"; "the instructor

makes the course practical"; "there is worthwhile and informative material in lectures that doesn't duplicate the text"; "the course has excellent content"; etc.

Nonstructured response

Ideal: —
 Important: (.50) S7
 Best: ASR = .49 (Range: .33-.60) S35, S6, S21, S45

Structured response

Ideal: ASR = .57 (Range: .44-.64) S12, S31, S13, S32
 Important: ASR = .43 (Range: .38-.58) S23, S36, S37, S24
 Best: (.62) S47

11. Nature and usefulness of supplementary materials and teaching aids: "the homework assignments and supplementary readings were helpful in understanding the course"; "the instructor made good use of teaching aids such as films and other audio-visual materials"; etc.

Nonstructured response

Ideal: —
 Important: (.93) S33
 Best: ASR = .63 (Range: .26-.86) S35, S6, S20

Structured response

Ideal: ASR = .49 (Range: .40-.71) S31, S13, S32, S12, S14
 Important: ASR = .74 (Range: .70-.77) S27, S9
 Best: (.44) S19

12. Difficulty of the course (and workload): "the instructor covered the right amount of material"; "the difficulty of the course was reasonable"; "the pace of the course was not too fast or too slow"; "the amount of outside readings seemed appropriate"; etc.

Nonstructured response

Ideal: (.40) S42
 Important: ASR = .69 (Range: .45-.93) S17, S33, S21, S35, S6
 Best: ASR = .74 (Range: .60-.83) S21, S35, S6

Structured response

Ideal: ASR = .73 (Range: .72-.76) S12, S13, S32, S31
 Important: ASR = .90 (Range: .86-.93) S23, S36, S16, S29
 Best: —

13. Instructor's fairness; impartiality of evaluation of students; quality of examinations: "grading in the course was fair"; "the instructor has definite standards and is impartial in grading"; "the exams reflect material emphasized in the course"; "test questions were clear"; "coverage of subject matter on exams was comprehensive"; etc.

Nonstructured response

Ideal: (.40) S42
 Important: ASR = .46 (Range: .11-.87) S2, S8, S41, S17, S7, S33
 Best: ASR = .49 (Range: .13-.88) S6, S45, S10, S20, S21, S35, S22

Structured response

Ideal: ASR = .43 (Range: .20-.70) S30, S38, S14, S12, S31, S32, S13, S30, S43

Important: ASR = .49 (Range: .05-1.00) S3, S4, S29, S26, S24, S18, S28, S16, S23, S15, S36, S37, S11, S27, S44, S9, S40, S34

Best: ASR = .75 (Range: .50-.90) S5, S47, S46

14. Classroom management: “the instructor controls class discussion to prevent rambling and confusion”; “the instructor maintained a classroom atmosphere conducive to learning”; “students are allowed to participate in deciding the course content”; “the teacher did not ‘rule with an iron hand’ ”; etc.

Nonstructured response

Ideal: —

Important: —

Best: ASR = .65 (Range: .28-.98) S21, S6, S20, S35

Structured response

Ideal: —

Important: ASR = .63 (Range: .42-.70) S3, S29, S4

Best: ASR = .70 (Range: .60-.80) S46, S25, S5

15. Nature, quality, and frequency of feedback to students: “the instructor gave satisfactory feedback on graded material”; “criticism of papers was helpful to students”; “the instructor told students when they had done a good job”; “the teacher is prompt in returning tests and assignments”; etc.

Nonstructured response

Ideal: —

Important: —

Best: ASR = .46 (Range: .13-.86) S6, S35, S20

Structured response

Ideal: ASR = .74 (Range: .60-.82) S12, S31, S32, S13, S14

Important: ASR = .72 (Range: .68-.76) S37, S23, S24

Best: —

16. Encouragement of questions and discussion; instructor’s openness to opinions of others: “students felt free to ask questions or express opinions”; “the instructor stimulated class discussions”; “the teacher encouraged students to express differences of opinions and to evaluate each other’s ideas”; “the instructor invited criticisms of his or her own ideas”; “the teacher appeared receptive to new ideas and the viewpoints of others”; etc.

Nonstructured response

Ideal: (.19) S42

Important: ASR = .44 (Range: .23-.68) S17, S2, S7, S33, S8

Best: ASR = .45 (Range: .04-.74) S35, S22, S6, S21, S20, S45, S10

Structured response

Ideal: ASR = .40 (Range: .24-.70) S32, S13, S12, S31, S30, S14, S39, S43, S38

Important: ASR = .53 (Range: .33-.65) S16, S37, S23, S11, S24, S44, S9, S26, S40, S4, S3, S36, S33

Best: ASR = .64 (Range: .60-.68) S5, S47, S46

17. Intellectual challenge; encouragement of independent thought: “this course challenged students intellectually”; “the teacher encouraged students to think out answers and follow up ideas”; “the teacher attempts to stimulate creativity”; etc.

Nonstructured response

Ideal: (.22) S42
 Important: (.79) S33
 Best: ASR = .60 (Range: .15-1.00) S45, S6, S22, S21, S48

Structured response

Ideal: ASR = .39 (Range: .10-.90) S39, S31, S32, S13, S12, S14, S38
 Important: ASR = .40 (Range: .20-.50) S44, S11, S40, S37, S23, S24, S34
 Best: (.44) S19

18. Instructor’s concern and respect for students; friendliness of instructor: “the instructor seems to have a genuine interest in and concern for students”; “the teacher took students seriously”; “the instructor established good rapport with students”; “the teacher was friendly toward all students”; etc.

Nonstructured response

Ideal: ASR = .28 (Range: .16-.39) S42, S49
 Important: ASR = .23 (Range: .06-.60) S2, S33, S41, S17, S7, S8
 Best: ASR = .27 (Range: .04-.48) S6, S22, S48, S20, S21, S1, S10, S45, S35

Structured Response

Ideal: ASR = .48 (Range: .22-.80) S13, S31, S12, S32, S14, S38, S39
 Important: ASR = .60 (Range: .25-1.00) S18, S9, S3, S11, S29, S24, S4, S16, S37, S23, S28, S26, S44, S15, S34, S36, S40
 Best: ASR = .46 (Range: .20-.64) S25, S46, S5, S47

19. Instructor’s availability and helpfulness: “the instructor was willing to help students having difficulty”; “the instructor is willing to give individual attention”; “the teacher was available for consultation”; “the teacher was accessible to students outside of class”; etc.

Nonstructured response

Ideal: —
 Important: ASR = .47 (Range: .13-.78) S33, S17, S7, S2
 Best: ASR = .20 (Range: .04-.33) S6, S20, S21, S22, S10, S1, S35

Structured response

Ideal: ASR = .45 (Range: .32-.53) S32, S13, S31, S30, S12, S14
 Important: ASR = .44 (Range: .26-.62) S25, S9, S3, S27, S37, S23
 Best: (.26) S46

TABLE II. Association Between College Students' Overall Evaluation of Their Teachers and Their Evaluation of the Specific Characteristics of These Teachers

Note: Following each ASR (overall average standardized rank, as explained in the text) is the range of standardized ranks and a listing of the studies on which the ASR is based. These studies are ordered from high to low with respect to their standardized ranks for the dimension under consideration.

S50 refers to study no. 50, S51 refers to study no. 51, and so on through S72 (study no. 72).

The studies are identified and briefly described in Appendix B.

The full label for each dimension (category of characteristics) and examples of responses can be found in Table I.

1. Stimulation of interest	ASR = .20 (Range: .10-.35)	S50, S61, S63, S60, S65, S70, S68, S57, S64, S56, S66, S58, S71, S69, S53
2. Enthusiasm	ASR = .46 (Range: .13-.80)	S66, S50, S68, S55, S62, S70, S57, S72, S65, S64, S63, S53, S60
3. Knowledge of subject	ASR = .48 (Range: .10-.70)	S53, S62, S71, S54, S56, S69, S63, S68, S55, S58, S66, S67, S72
4. Intellectual expansiveness	ASR = .54 (Range: .39-.77)	S58, S72, S56
5. Preparation and organization	ASR = .41 (Range: .09-.94)	S68, S58, S51, S66, S50, S72, S61, S71, S62, S59, S52, S55, S69, S63, S54
6. Clarity and understandableness	ASR = .25 (Range: .04-.72)	S52, S66, S71, S65, S70, S55, S69, S54, S60, S59, S67, S62, S64, S50, S61, S53, S57, S51
7. Elocutionary skills	ASR = .49 (Range: .08-.95)	S59, S62, S52, S61, S68, S58
8. Sensitivity to class level and progress	ASR = .40 (Range: .13-.84)	S71, S62, S52, S50, S51, S54 S53
9. Clarity of objectives and requirements	ASR = .45 (Range: .08-1.00)	S58, S62, S67, S71, S68, S72, S50, S69
10. Value of course material	ASR = .70 (Range: .21-.92)	S66, S54, S59, S62, S51, S53, S71, S56, S50
11. Usefulness of supplementary materials	ASR = .72 (Range: .13-1.00)	S51, S58, S68, S62, S66, S67
12. Difficulty (workload)	ASR = .87 (Range: .54-1.00)	S53, S67, S57, S59, S62, S72
13. Fairness and evaluation	ASR = .72 (Range: .15-1.00)	S72, S66, S52, S64, S53, S69, S67, S62, S50, S59, S71, S60, S70, S58, S63, S65, S68, S51, S55, S54
14. Classroom management	ASR = .65 (Range: .26-.96)	S52, S62, S69, S71, S72, S53
15. Feedback to students	ASR = .87 (Range: .79-.96)	S58, S62, S54, S66, S50
16. Encouragement of discussion (openness)	ASR = .60 (Range: .09-.90)	S67, S52, S62, S53, S51, S60, S58, S72, S66, S50, S56, S59, S61, S54, S64, S65, S69, S70
17. Intellectual challenge	ASR = .39 (Range: .10-.94)	S64, S60, S65, S70, S50, S58, S59, S51, S54, S66, S67, S53, S51, S61
18. Respect for students (friendliness)	ASR = .65 (Range: .25-1.00)	S71, S56, S65, S58, S50, S59, S60, S52, S70, S55, S62, S72, S67, S64, S63
19. Availability and helpfulness	ASR = .74 (Range: .53-.94)	S58, S68, S67, S69, S66, S71, S50, S61

TABLE III. Characteristics Most Associated with Superior Teachers and Superior Teaching, as Determined by Various Procedures

Note: For this table, the standardized ranks in Tables I and II have themselves been ranked, with the dimensions or categories of the highest ranking given here (see the numbers not in parentheses). The overall average standardized rank is based on only one study, it and its rank placement in this table are given in brackets. The three subdivisions within the nonstructured-response set of studies (cols. 1-3) and within the structured-response set (cols. 4-6) have been considered together (see text); for each of these sets, the dimensions have been roughly ordered and clustered, with results based on only one study (in brackets) given a little less "weight." Conceivably, categories no. 10 (value of course material) and no. 15 (feedback to students) belong among the highest dimensions within the nonstructured-response set of studies, but they have been excluded because the results for each are based on too few studies across the three subdivisions of studies (see Table I). For the same reason, category no. 9 (clarity of course objectives and requirements) has been excluded from the results shown for the structured-response set of studies. The full label for each dimension and examples of responses can be found in Table I.

	Nonstructured-response studies			Structured-response studies			Correlational studies (overall evaluation X specific evaluations) (7)
	Ideal (1)	Important (2)	Best (3)	Ideal (4)	Important (5)	Best (6)	
Respect stdnts	6.5 (.28)	1 (.23)	2 (.27)	Knowledge	2.5 (.18)	3 (.26)	1 (.11)
Knowledge	5 (.23)	2 (.24)	3.5 (.33)	Interesting	1 (.14)	4 (.28)	2.5 (.26)
Interesting	6.5 (.28)	3 (.36)	3.5 (.33)	Class progress	4 (.19)	1 (.20)	5 (.40)
Availability	---	7 (.47)	1 (.20)	Clarity	5 (.21)	2 (.22)	6.5 (.42)
Discussion	[2 (.19)]	4 (.44)	7 (.45)	Enthusiasm	2.5 (.18)	6.5 (.35)	6.5 (.42)
Clarity	1 (.17)	10.5 (.60)	5 (.39)	Organization	6 (.29)	5 (.29)	11 (.48)
Enthusiasm	[3.5 (.22)	9 (.53)	11 (.50)	Challenge	7 (.39)	8 (.40)	[8.5 (.44)]
Fairness	[9.5 (.40)]	6 (.46)	9.5 (.49)	Availability	10 (.45)	10 (.44)	[2.5 (.26)]
Organization	11 (.56)	12 (.63)	6 (.44)	Discussion	8 (.40)	12 (.53)	14 (.64)
Elocution	[12.5 (.74)]	5 (.45)	12 (.52)	Respect stdnts	11 (.48)	13 (.60)	10 (.46)

The highest ranked dimensions for the nonstructured-response set of studies are the instructor's concern or respect for students (including friendliness) and the instructor's knowledge of subject matter. Next in importance are the instructor's stimulation of students' interest, the instructor's availability and helpfulness, the instructor's encouragement of questions and discussion (including openness to others' opinions), and the instructor's ability to explain clearly. Lower in rank order are the instructor's enthusiasm for the subject or for teaching, and the instructor's impartiality, followed by the instructor's preparation for (and organization of) the course, and the instructor's elocutionary skills.

The results for the studies based on the structured responses of students show similarities as well as some interesting differences with the results of the nonstructured-response set of studies. For the structured-response set, the dimensions of knowledge and of stimulation of interest again are highly ranked by students (perhaps even a bit more highly so than when student responses are not structured). Again placing in the list of the ten most important dimensions are clarity of explanation, enthusiasm, and preparation-organization; if anything, each of these three categories of characteristics appears to be a little more important than it was in the structured-response set of studies. The most dramatic difference between the two sets of rank orderings is the difference in the relative importance of the dimension of instructor's concern and respect for students. This dimension, which ranks as one of the two highest in the top-ten list based on the nonstructured-response set of studies, ranks as one of the two lowest in the top-ten list based on the structured-response set of studies. Similar to this result, although a little less dramatic, is the shift downward in rank of two other interpersonally focused dimensions—namely, instructor's availability (helpfulness) and instructor's encouragement of class questions and discussion (openness to the opinions of others). These two dimensions are near the bottom of the top-ten list based on the structured responses of students, whereas they are among the five most important dimensions when students' responses are not structured. Finally, neither the instructor's impartiality nor elocutionary competence is still in the top-ten ranks for the nonstructured-response set of studies; these dimensions have been displaced by the instructor's sensitivity to class level and progress and by his or her ability to challenge students intellectually.

As can be seen in Table I, for each dimension there is a range of results. Such differences are probably due, in part, to differences in the year (or historical period) in which the research was done, the type of school in which the research was conducted, and the like. Even within a given study and a given sample of students, of course, there is some variation as to

what characteristics are felt to be most important to good teaching either by students considered individually (e.g., see Permut, 1973) or by subgroups of students. The subgroups of students most often compared within a study have been those obtained by dividing students in a sample by their sex (Bousfield, 1940; Crawford and Bradshaw, 1968; Gadzella, 1968a; Grasha, 1975; Kreuger, 1936; Lehmann, 1966; Levinthal, 1974; Morton, 1965; Mueller, Roach, and Malone, 1971; Pogue, 1967; Riley et al., 1950; and Taylor, 1959), by their class in college (Brewer and Brewer, 1970; Crouch and Leathers, 1951; Gadzella, 1967, 1968a; Geyer, 1946; Morton, 1965; Pogue, 1967; Remmers, as cited in Smith, 1944 and in Geyer, 1946; Riley et al., 1950; Smith, 1944; Turner, Evans, Hale, Cairns, and Maleski, 1969; and Yourglic, 1955), and by academic field or division (Birney et al. 1960; Gadzella, 1967, 1968b; Hussain and Leestamper, 1968; Levinthal, 1974; Mueller et al., 1971; Musella and Rusch, 1968; Riley et al., 1950).³ Although such subgroups, in general, are rather highly similar with respect to the characteristics that are typically preferred in teachers, certain differences do appear in the studies.

In the research by Bousfield (1940), Gadzella (1968a), Morton (1965), Pogue (1967), and Riley et al. (1950), some differences were found between male and female students in the relative emphasis they placed on various attitudinal and behavioral characteristics of college teachers, although these differences were usually not very large and the results are not consistent across studies. Moreover, little, if any, difference in preferences was found between these two groups in other studies (Crawford and Bradshaw, 1968; Gadzella, 1967, 1968b; Grasha, 1975; Lehmann, 1966; and Mueller et al., 1971).

The similarity among students in different college classes is also generally high, although certain differences do appear. The most consistent of them across studies is the tendency for underclassmen (either freshmen alone or freshmen and sophomores combined) to place somewhat more emphasis than do upperclassmen on the importance of the instructor's fairness or impartiality as well as on his or her ability to get along with students (see especially Brewer and Brewer, 1970; Crouch and Leathers, 1951; Gadzella, 1967; Remmers, as cited in Smith, 1944; Riley et al., 1950; and Smith, 1944). Moreover, the research suggests that students at higher college-class levels place somewhat greater value than do those at lower levels on the organization of the subject matter or the course and on the instructor's enthusiasm for teaching (see especially Brewer and Brewer, 1970; Crouch and Leathers, 1951; and Pogue, 1967).

Certain consistent differences among academic fields also appear in the studies in this area. Students majoring in the fields of natural sciences, physical sciences, and mathematics, as well as students describing their

preferences for teachers in these fields (whether or not these fields were their majors), put relatively more stress than do other students on the importance of teachers being able to explain clearly (see especially Birney et al., 1960; Gadzella, 1968b; Musella and Rusch, 1968; and Riley et al., 1950). Moreover, there appears to be greater emphasis in these fields, as well as in the social sciences, on the instructor's preparation and organization of course material (see especially Birney et al., 1960; Musella and Rusch, 1968; and Riley et al., 1950). Finally, emphasis on the ability of the teacher to encourage thought and to be intellectually challenging is more evident in the social sciences, humanities, and fine arts than in other fields (see especially Musella and Rusch, 1968; and Riley et al., 1950).

CHARACTERISTICS ASSOCIATED WITH STUDENTS' OVERALL EVALUATION OF THEIR INSTRUCTORS

Thus far, information has been presented regarding students' reported preferences for various characteristics in college teachers, not about the actual importance of these preferences in students' evaluations of their teachers. It is of more than routine interest to compare the attitudes, behaviors, and teaching practices that students say are important to good teaching with the actual importance of these characteristics to students' overall assessment of their specific teachers. On the plausible assumption that each student's global evaluation of an instructor is an additive combination of the student's evaluations of specific aspects of teachers or their teaching behaviors, weighted by the student's estimation of the relative importance of these aspects to good teaching, it would be expected that students' overall evaluations of instructors would be more highly associated with characteristics that students generally consider to be more important to good teaching than those they consider to be less so (see Crittenden and Norr, 1973). That this indeed is empirically true is not inevitable, however.

It is altogether possible that a trait considered by students to be highly important to good teaching (or highly characteristic of ideal or best teachers) does not particularly differentiate the better from the poorer teacher in actual teacher-rating situations. For example, excellent knowledge of the subject matter may be especially characteristic of students' best teachers, but this does not necessarily mean that their less than best teachers are any the less knowledgeable. These lower-rated teachers, too, may be highly expert in their subject; if so, this would simply mean that the dimension of subject-matter knowledge would be unlikely to discriminate among teachers with respect to their overall ratings on teacher evaluation forms. Furthermore, the differential weights of various

teacher characteristics that students, in effect, say they use (or will use) in evaluating their teachers—determined by asking them about their ideal or best teachers or about the importance of various characteristics to good teaching—are not necessarily the weights they actually use, in practice, when forming and reporting a global impression of each of their current teachers (Permut, 1973).

A number of studies were located containing correlations between students' overall evaluations of their instructors and their ratings of various specific attitudinal or behavioral characteristics of these instructors (see Appendix B). Using the same coding and standardizing procedures as those used in the earlier analysis of student preferences, overall average standardized ranks were calculated for the same teaching dimensions as those in the prior analysis.⁴ Results are shown in Table II. The ten highest ranking dimensions are given in Table III (col. 7), where comparisons can be made with the results of the earlier analysis (cols. 1-6). Before discussing these comparisons, it should be emphasized that the studies of overall evaluation of teachers involve different samples of students than the studies in which students described the characteristics they wanted in their teachers and thought to be important to good teaching. To some extent, then, these sampling differences may account for any differences that are found in the results among the studies. (For a study in which the same sample of students rated characteristics in terms of their importance for good teaching and also rated their instructors both on an overall evaluation item and a set of specific evaluation items, see Crittenden and Norr, 1973.)

Ranking the characteristics in terms of the degree to which they correlate with the overall evaluation of instructor produces results more similar to the results in the structured-response set of studies (cols. 4-6) than to those in the nonstructured-response set (cols. 1-3), although there are certain differences with both sets. (As in the earlier analysis, the three subdivisions of either of these two sets of studies are considered together.) The dimensions having the five highest correlations with actual overall evaluation are the following: stimulation of interest, clarity of explanation, intellectual challenge, sensitivity to class level and progress, and preparation or organization.⁵ These five are also in the top-ten list of the structured-response set of studies, but only three of them are so placed in the nonstructured-response set. Moreover, each of these five is higher in rank in the list of the correlates of overall evaluation of the instructor than it is in either the structured-response list of preferred characteristics (excepting sensitivity to class level and progress) or the nonstructured-response list. (Note that the differences in rank placement are smaller for the comparison with the structured-response list than they are for the

comparison with the nonstructured-response list.) In comparison to the nonstructured-response and structured-response lists, knowledge of subject matter is now of rather lower importance. Also of less importance than previously are the three interpersonally focused dimensions of instructor concern for students, helpfulness, and encouragement of class discussion, each having slipped out of top-ten placement altogether. Finally, two new dimensions, clarity of course objectives and requirements, and instructor's intellectual expansiveness (intelligence), now appear in the top-ten list.⁶

It may be noted at this point that the ratings of specific instructor characteristics take different forms. Four of these follow, each regarding the instructor's organization of course material:

1. How satisfied were you with this instructor's organization of the course material: (a) highly satisfied, (b) somewhat satisfied, (c) somewhat dissatisfied, (d) highly dissatisfied.

2. Compared to other instructors you have known, how would you rate your instructor as to his or her organization of the course material: (a) one of the best, (b) somewhat above average, (c) about average, (d) somewhat below average, (e) one of the worst.

3. To what degree was the course material organized by your instructor: (a) highly organized, (b) somewhat organized, (c) somewhat disorganized, (d) highly disorganized.

4. How applicable to your instructor is this statement, "The instructor made certain that the course material was well organized": (a) highly applicable, (b) somewhat applicable, (c) somewhat inapplicable, (d) highly inapplicable.

In the first two cases, each student actually evaluates the teacher in a specific area; the student's judgment presumably is based on the degree to which the perceived attitude or behavior of the teacher matches the student's preference in this area. By contrast, in the last two cases, the student is merely asked to describe a characteristic of the instructor or to indicate the extent to which a given description holds true. Here, the student makes explicit what he or she has observed, but does not directly or explicitly offer an evaluation. The student's own attitude or judgment must be inferred (cf. Levinthal et al., 1971). For example, a student's description of a teacher as highly organized is taken to mean that the student highly approves of the instructor's performance in this particular area. However, this inference is accurate only if the student values high organization of a course, not, as it might happen, if that student prefers a teacher to be only moderately organized or perhaps even somewhat disorganized. Therefore, it would be expected that the degree of congruence between the characteristics of a teacher, as described by the student, and

the student's own preference would be more highly predictive of the student's overall evaluation of the teacher than would the student's description alone. Surprisingly, the few relevant studies that have been done in the area offer almost no support for this expectation (in Costin and Grush, 1973, compare the data in Table I with those in Table IV; in Hall, 1970, see the first column in Table 1; and in Levinthal, 1974, compare "D" scores with "O" scores in Appendix E; but see Gagné and Allaire (1974) for a discussion of how this lack of support may be due to certain methodological problems)⁷

FACTOR ANALYSIS OF TEACHER EVALUATION QUESTIONNAIRES

In addition to the relationships between the overall rating of the instructor (or the course) and the ratings of specific instructor characteristics, the interrelationships among the specific ratings themselves are of interest. Comparatively few studies present the full array of interassociations among specific ratings, but a large number of studies have reported the results of factor analyses based on such intercorrelations. For each of nearly five dozen pieces of research containing appropriate and sufficient data on a factor analysis (or a closely similar analysis),⁸ the highest loading items on each factor reported in the study were coded into the categories used throughout the present analysis. Added to these were a category for items concerned with the outcome of instruction (see footnote 5) and a category for items dealing with the overall evaluation of the teacher or the course.⁹

For the set of factor-analytic studies under consideration, there are instances in which the highest loading items on a factor could be coded into the same category or dimension (with all other items of the evaluation questionnaire having lower loadings on this factor). For example, in a study by Arreola (1973) of students at Florida State University, the four highest loading items (of 21 items) on a factor entitled Course Organization were the following: the course was well organized; the instructor appeared to relate the course concepts in a systematic manner; the direction of the course was adequately outlined; the instructor's class presentation made for easy note taking. This particular factor in this study, then, turns out to be a relatively "pure" indicator of category no. 5 of the present analysis. However, the exclusive representation of specific evaluation items from the same category of characteristics on a consequently pure factor is far from the typical occurrence in the studies under consideration. In general, the highest loading items on any given factor comes from more than one of the dimensions of the present analysis, although often one, two, or three of these dimensions predominate among the highest loading items.

Systematic track was kept of which characteristics representing which dimensions were the highest loading items in each factor in each of the factor-analytic studies. Figure 1 shows the overall pattern of the results of this procedure. The dimensions that often combine with one another as the highest loadings on a factor are connected by a line, with the heavier line reserved for those characteristics that most often are combined with each other. Despite the profusion of connections in this figure, a fairly consistent and interpretively meaningful pattern does emerge; indeed, this pattern supports the view of Widlak et al. (1973) that instructors primarily enact three different roles, each of which has certain distinctive teaching tasks associated with it.

On the left-hand side of the figure, there are six categories of items, any two or more of which frequently load highly on the same factor: overall evaluation of the teacher or the course (no. OE, see footnote 9); instructor's stimulation of interest (no. 1), instructor's clarity and understandableness (no. 6); instructor's preparation and organization of the course (no. 5); instructor's enthusiasm (no. 2); and instructor's knowledge of subject matter (no. 3). (Of these six dimension, the first four are most frequently connected.) Connected to at least two of these dimensions, but not particularly to each other, are outcome of instruction (no. IO, see footnote 5), instructor's intellectual expansiveness (no. 4), instructor's elocutionary skill (no. 7), and instructor's sensitivity to class level and progress (no. 8). Most generally, these ten dimensions represent the kinds of attitudes, behaviors, and teaching practices found during an instructor's *presentation* of material. As Widlak et al. (1973) have put it, this particular teaching task is part of the instructor's role as Actor or Communicator.

It may be seen in this first cluster of dimensions that items measuring overall evaluation of the teacher or the course are connected with stimulation of interest, enthusiasm, knowledge of subject matter, preparation and organization of material, clarity, and instructional outcome for the student. These very dimensions are among the categories that showed the highest correlations with the overall evaluation of the teacher in the analysis presented earlier (see Table II, Table III, and footnote 5). The factor-analytic studies analyzed in this section are generally not the same studies as those in which specific evaluative items are correlated with overall assessment of the teacher. Since different samples of students are consequently involved in these two sets of studies, the consistency in results between the two sets increases confidence in the generality of these particular findings.¹⁰

A second cluster of characteristics (placed on the right-hand side of the figure) primarily involves the teaching task of *facilitation* that is associated with the instructor's role as Interactor or Reciprocator (Widlak et al., 1973).

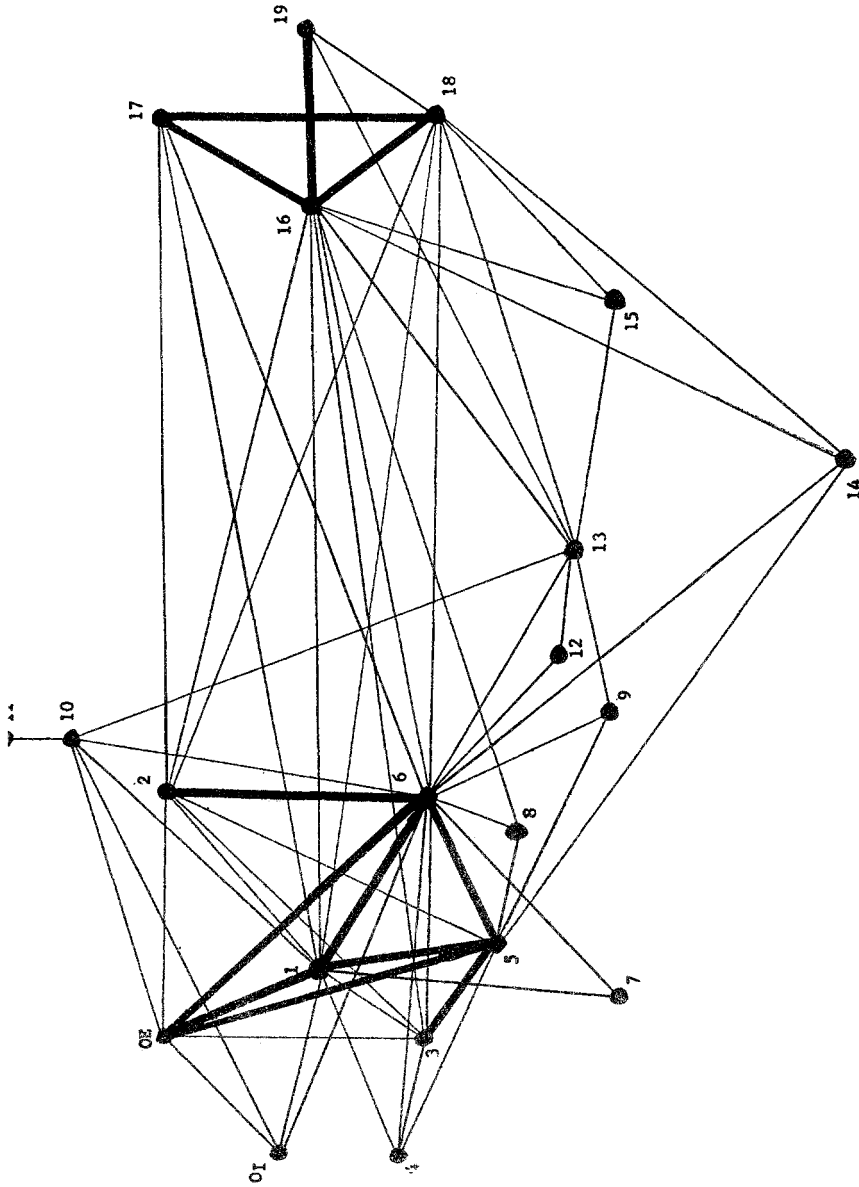


Fig. 1. Most frequently "connected" categories of teacher and instructional characteristics in factor-analytic studies (as explained in the text). Nos. 1-9 refer to the categories of characteristics (dimensions) used throughout the analysis, as given in Table I. OI refers to the dimensions of "outcome of instruction" (see footnote 5); OE refers to the dimension of "overall evaluation" (see footnote 9).

This cluster comprises the categories of instructor's friendliness and concern or respect for students (no. 18), instructor's openness to others' opinions, including encouragement of class questions and discussion (no. 16), instructor's intellectual challenge and encouragement of independent thought (no. 17), and, to a lesser extent, instructor's availability and helpfulness (no. 19). When any one of the first three of these four dimensions combine with elements of the first cluster of dimensions (as the highest loading items on a factor), it does so most frequently with instructor's stimulation of interest (no. 1), instructor's clarity and understandableness (no. 6), and/or instructor's enthusiasm (no. 2).

The remaining set of characteristics have been placed between those in the two clusters just described, for they have direct or indirect connections with characteristics in both clusters. Thus, classroom management (no. 14) has a connection with preparation-organization (no. 5) and clarity-understandableness (no. 6) in the first cluster and with encouragement of questions or openness to others' opinions (no. 16) and concern for students or friendliness (no. 18) in the second. With one exception, the dimension constituted by the instructor's fairness, impartiality of evaluation, and quality of examinations (no. 13) is connected with these very same characteristics. Furthermore, this fairness-of-evaluation dimension is also connected with the following dimensions, each of which, in turn, is connected with dimensions in one or the other of the two clusters: nature and value of the course material (no. 10), clarity of course objectives and requirements (no. 9), difficulty of course and workload (no. 12), and the quality and frequency of feedback to students (no. 15). From a general perspective, these various dimensions in the middle of the figure can be considered to be aspects of the instructor's role as Director or Administrator, as associated with the teaching task of *regulation* (Widlak et al., 1973).

To conclude this section, certain methodological considerations can be noted. Item intercorrelations (on which factor analyses are based) are subject to various systematic artifacts and "errors," including halo effect, central tendency, proximity error, and logical error (see especially Guilford, 1954, ch. 11; and Doyle, 1975, ch. 3). Such errors influence the strength and patterning of the correlations among specific items and thus partially determine the factors that emerge in a factor analysis. Factor-analytic results are also dependent on the nature of the items that are put into the factor analysis and of the sample(s) of individuals who have supplied the information. For instance, it is not possible to find that items dealing with instructors' organization of the course and with their ability to explain clearly load highly on the same factor, if these items are not on the teacher evaluation questionnaire in the first place. Similarly, a factor

analysis done on information supplied by a highly unrepresentative sample of students may produce factors unlike those to be found in more typical student populations. Another set of problems is produced by the technical "indeterminancies" and interpretive difficulties that are involved in any factor analysis (Harman, 1967). That all this is the case probably argues against placing too much importance on any one factor analysis and supports the effort to compare results across studies. Even so, it should be kept in mind that studies may not always be fully comparable and that multistudy comparison cannot really resolve any ambiguities inherent in the factor solutions themselves.

CONCLUDING COMMENTS

The attitudes, behaviors, and pedagogical practices that are most associated with superior college teaching have been determined in four ways in this analysis: (1) characteristics that students report as being most associated with ideal or best teachers and as most important to effective teaching, with students furnishing lists of characteristics of their own choosing; (2) characteristics that students report as being most associated with ideal or best teachers and as most important to effective teaching, with students responding to pre-set lists of characteristics; (3) specific items on teacher evaluation questionnaires that are most strongly associated with the global evaluation of the instructor; and (4) specific evaluation items that most frequently combine with global evaluation items to form the highest loadings on the same factor in factor-analytic studies.

Across studies using any of these types of measurement, stimulation of interest and clarity (understandableness) are the two dimensions that are most consistently highly associated with superior college teachers or teaching. Knowledge of subject matter is also highly characteristic of the superior teacher or effective teaching, although somewhat less consistently so across studies. The two other consistently important characteristics are those that have been classified as instructor's preparation for (and organization of) the class and instructor's enthusiasm for the subject matter or for teaching. Furthermore, characteristics codable into one or another of these five categories are the very ones that frequently load highly with one another (in combinations of two or more) on the same factor in factor-analytic studies (see Figure 1); they form the core of a set of instructor attitudes, behaviors, and practices that constitute the teaching task of subject-matter presentation (instructor in the role of Actor or Communicator).

Another set of characteristics primarily comprises the regulative activities of the instructor in his or her role of Director or Administrator:

classroom management; impartiality of evaluation; clarity of course objectives and requirements; difficulty of the course and workload; frequency and quality of feedback to students; and the like. Relative to other characteristics, these are rather consistently lower in importance for superior teaching, at least by the several indicators used in the present analysis.

Friendliness (concern and respect for students), helpfulness (availability), and openness to others' opinions (encouragement of class questions and discussion), dimensions primarily involving the teaching task of facilitation (instructor in the role of Interactor or Reciprocator), are among the most frequently mentioned characteristics when students freely describe their ideal or best teachers and the characteristics they see as important to good teaching. However, when students respond to a preset list of characteristics, these three dimensions are of less importance. Why this should be so is not clear. One speculation is that, although these sorts of social-emotional matters may indeed be a central aspect in persons' (including students') general consideration of others and a central part of their response repertoires, they may become less important as more specific and structured situations increase the saliency of other considerations (cf. Newcomb et al., 1965, ch. 2 and 3; Brown, 1965, ch. 6; and Tagiuri, 1969). At any rate, compared to both the unstructured and structured preferences of students, these characteristics rank even lower in importance (below the top-ten ranks) in terms of the strength of association between students' ratings of their instructors on each of them on teacher evaluation questionnaires and their overall ratings of these instructors. Moreover, these characteristics are not particularly likely to combine with global evaluation items to form distinctive factors in factor-analytic studies.

It is sometimes supposed that students most want to take courses with teachers who are warm, friendly, helpful, open to discussion with them, and generally respectful of them and their opinions. It is also sometimes said that consideration of such matters unduly influences the overall evaluation of instructors during formal teacher evaluations. Neither of these suppositions is supported by the present analysis of existing research. It is true that characteristics of interpersonal facilitation are mentioned frequently when students respond freely about their preferences, but they are hardly the only characteristics to receive frequent mention. Moreover, as noted, they are typically less important than certain others when students judge the importance of predetermined descriptions of teachers and teaching. Of greater significance, these variables are less strongly associated with overall evaluation of actual instructors than are the instructor's stimulation of interest, clarity or understandableness, knowledge of subject matter, preparation and organization, and

enthusiasm for the subject and for teaching. That this latter set of variables appears to be of greater importance in the actual assessment of instructors (in addition to the fact that students explicitly report that these are among the characteristics they most want in a teacher) should be encouraging to those educators who argue for the meaningfulness of college students' opinions about teachers and teaching.

Still in all, conclusions should not be drawn too quickly or too easily about the significance, accuracy, and validity of college students' general views on teaching (in the abstract) as well as their more concrete assessments of actual teachers. Certain problems and issues remain. It is important to raise some of them briefly, although extended analysis will not be undertaken at this point.

To begin with, it should be emphasized that students' views of the ideal teacher, their listing of traits most important to good teaching, and their specification of the characteristics of their best teachers are essentially descriptions rather than explanations. Furthermore, as was discussed, descriptively important characteristics are not necessarily the characteristics that differentiate the more highly rated teachers from the lower rated ones on teacher evaluation questionnaires. Even when they are—that is, even when these descriptively important characteristics are among the characteristics that are most highly related to the overall evaluation of the teacher (or the course)—causal implications may still be unclear or ambiguous. Superiority on specific items of a teacher evaluation questionnaire does not necessarily explain the teacher's overall superiority (as perceived by students); for that matter, presumed superiority in specific areas of teaching may not even have produced the high overall rating. It could be that once students have an initial impression that a teacher is superior—an impression based, perhaps, on only minimal if not somewhat irrelevant cues—then the student, in turn, may come to see the teacher as especially competent in specific aspects of teaching, say, organization of the course, ability to explain clearly, knowledge of subject matter, enthusiasm for teaching the subject, or the like. (This phenomenon is discussed under the more general rubric of halo effect, the tendency to rate in terms of overall general impression without differentiating specific aspects, of allowing our total reaction to the person to color our judgment of each specific trait, Thorndike and Hagen, 1969, p. 432; for discussions and analyses of the halo effect in the context of college teacher evaluation, see Centra, 1973; Creager, 1950; Doyle, 1975, pp. 41-42; French-Lazovik, 1974; Hodgson, 1958; Hoyt, 1973; Miklich, 1969; Purohit and Magoon, 1971; Sockloff and Deabler, 1971; and Widlak et al., 1973). Conceivably, overall general impressions and judgments of specific traits of college professors may feed into one another and causally intertwine in complex ways.

Even assuming that college students build up their overall impressions

of their teachers from their independent judgments about specific characteristics of their teachers, the question can still be (and has been) raised as to whether students are in a position to make accurate judgments about certain matters, including the instructor's degree of knowledge of the subject matter of the course, the instructor's preparation and organization of the course, and the instructor's ability to explain clearly (see Brickman, 1966; Centra, 1966; Dansereau, 1973; Deegan, 1972; Diener, 1973; Gustad, 1967; Marshall, 1952; and Withington, 1944). The charge of possible judgmental inaccuracies is not applicable in all areas, of course. For example, if students evaluate a teacher as being highly interesting or stimulating to them, it cannot really be argued that the teacher is not. However, even in this case, it is still possible that, under certain circumstances, the teacher may be interesting and stimulating for presumably wrong or inauthentic reasons, for example, reasons of authoritative style, showmanship, and wit at the expense of substance and meaning (cf. Naftulin et al., 1973).

A note of caution about the present analysis itself may be added here. The attempt throughout has been to generalize from as much information as possible, to discover that which is common among studies, and to condense, systematize and summarize results of extant research. This effort to integrate the results of studies involves certain difficulties. The major problem hinges on the uncertainty of generalization when the comparability of separate pieces of research may be open to question and when the degree of representativeness of the samples of studies is not precisely known (see Feldman, 1971). Even apart from the issue of generalizability, procedures used in integrations often introduce certain distortions. In the present analysis, a wide array of attitudinal and behavioral characteristics of teachers have been classified into a smaller number of categories, and results of studies have been standardized and averaged. Although such condensing and averaging is probably desirable in general, these procedures can underplay the occurrence and importance of individual variation and atypical findings. For example, although it is important to know that more students in a particular study, say, select the category of stimulation of interest than any other category as most characteristic of the ideal teacher, it is also important to know something about individual variation in results (the proportion of students in the sample who did not feel this particular characteristic to be the most important, the characteristics that these students did choose, and so forth). Similarly, a characteristic can be typically important across studies, but knowing this should not prompt the dismissal of studies where it is not, nor should it stop the search for understanding the circumstances that might have generated atypical results.

For all the problems of integrations and reviews, they are important if

headway is to be made in an area of research. The present analysis, along with recent reviews of the research on teacher evaluation (Costin et al., 1971; Doyle, 1975; Dwyer, 1968; and Kulik and McKeachie, 1975), should help to provide a framework within which specific studies can be placed and to suggest future research that is likely to have the most payoff.

ACKNOWLEDGMENTS

This is the second paper of a projected series on teacher and course evaluation (see Feldman, 1976). I would again like to acknowledge the helpfulness of James L. Bess, who alerted me to materials that I may not otherwise have located.

REFERENCES

- Albino, J.E., and Liberty, P.G. (n.d.). Some methodological considerations in the use of factor analysis to determine dimensions underlying student ratings of courses and instructors. Unpublished manuscript.
- Aleamoni, L.M., and Spencer, R. E. (1973). The Illinois Course Evaluation Questionnaire: A description of its development and a report of some of its results. *Educational and Psychological Measurement* 33:669-684.
- Aleamoni, L. M., and Yimer, M. (1973). An investigation of the relationship between colleague rating, student rating, research productivity, and academic rank in rating instructional effectiveness. *Journal of Educational Psychology* 64:274-277.
- Apt, M. H. (1966). A measurement of college instructor behavior. Unpublished doctoral dissertation, University of Pittsburgh.
- Arreola, R. A. (1973). A cross-institutional factor structure replication of the Michigan State University SIRS faculty evaluation model. *College Student Journal* 7: 38-42.
- Asher, E. J., Jr. (1970). Differential weighting of teacher evaluation criteria. In *The Challenge and response of Institutional: Proceedings of the Ninth Annual Forum on Institutional Research*. (C. Fincher, ed.). The Association for Institutional Research.
- Axelrod, J. (1973). *The University Teacher as Artist: Toward an Aesthetics of Teaching With Emphasis on the Humanities*. San Francisco: Jossey-Bass.
- Bausell, R. B., and Magoon, J. (1972). The validation of student ratings of instruction: An institutional research model. Newark, Del.: College of Education, University of Delaware.
- Bejar, I. L., and Doyle, K. O. (1974). Generalizability of factor structures underlying student ratings of instruction. Paper read at the annual meeting of the American Educational Research Association.
- Bendig, A. W. (1954). A factor analysis of student ratings of psychology instructors on the Purdue Scale. *Journal of Educational Psychology* 45:385-393.
- Birney, R. C., Coplin, H. R., and Grose, R. F. (1960). *The class of 1959 at Amherst College: A preliminary report*. Amherst, Mass.: Committee on Guidance and Counseling, Amherst College.

- Blai, B., Jr. (1974). Viewpoints—to each his own! A parallel-perceptions inquiry. Bryn Mawr, Penn.: Office of Research, Harcum Junior College.
- Bogardus, E. S. (1946). Behavior patterns of college teachers. *Sociology and Social Research* 30:484-490.
- Bousfield, W. A. (1940). Students' ratings of qualities considered desirable in college professors. *School and Society* 51:253-256.
- Brandenburg, D. C. (1975). The general concept of validity applied to student ratings or please, General Custer, what are we doing here? Paper read at the annual meeting of the American Educational Research Association.
- Brandis, R. (1964). The rehabilitation of university undergraduate teaching. *Educational Record* 45:56-63.
- Breed, F. S. (1927). Factors contributing to success in college teaching. *Journal of Educational Research* 16:247-253.
- Brewer, R. E., and Brewer, M. B. (1970). Relative importance of ten qualities for college teaching determined by pair comparisons. *Journal of Educational Research* 63:243-246.
- Brickman, W. W. (1966). Student evaluation of professors. *School and Society* 34:143-144.
- Bridges, C. M., Ware, W. B., Brown, B. B., and Greenwood, G. (1971). Characteristics of best and worst college teachers. *Science Education* 55:545-553.
- Brooks, T. E., Tarver, D. A., Kelley, H. P., Liberty, P. G., Jr., and Dickerson, A. D. (1971). Dimensions underlying student ratings of courses and instructors at the University of Texas at Austin: Instructor Evaluation Form 2. Research Bulletin RB-71-4. Austin, Texas: Measurement and Evaluation Center, University of Texas at Austin.
- Brown, R. (1965). "Social Psychology." New York: Free Press.
- Caffrey, B. (1969). Lack of bias in student evaluations of teachers. Proceedings of the 77th Annual Convention of the American Psychological Association 4:641-642.
- Caffrey, B. (1973). Use of multiple regression analysis and factor analysis to design a brief, objective assessment of instructor performance. Paper read at the annual meeting of the Southeastern Psychological Association.
- Carter, R. E. (1968). The effect of student characteristics on three student evaluations of university instruction. Unpublished doctoral dissertation, Indiana University.
- Carver, D. W., and Liberty, P. G. (1973). A comparison of two approaches in factor studies of student ratings of courses and instructors. Paper read at the annual meeting of the American Educational Research Association.
- Case, H. W. (1952). Student evaluation of teaching techniques. *Journal of Engineering Education* 42:446-450.
- Cassel, R. N., (1971). A student course evaluation questionnaire. *Improving College and University Teaching* 19:204-206.
- Centra, J. A. (1966). Problems of evaluating teaching. In "Current Issues in Higher Education." Washington, D.C.: Association for Higher Education.
- Centra, J. A. (1972). Strategies for improving college teaching. ERIC Clearinghouse on Higher Education report no. 8 Washington, D.C.: American Association for Higher Education.
- Centra, J. A. (1973). The Student Instructional Report: Comparisons with alumni ratings; item reliabilities; the factor structures. SIR report no. 3. Princeton, N.J.: Educational Testing Service.

- Centra, J. A. (1975). Colleagues as raters of classroom instruction. *Journal of Higher Education* 46:327-337.
- Centra, J. A., and Linn, R. L. (1973). Student points of view in ratings of college instruction. *Research Bulletin RB-73-60*. Princeton, N.J.: Educational Testing Service.
- Champlin, C. D. (1928). The preferred college professor. *School and Society* 27:175-177.
- Champlin, C.D. (1953). Professors of education. *Education* 73:346-350.
- Clinton, R. J. (1930). Qualities college students desire in college instructors. *School and Society* 32:702.
- Cloer, H. A. (1970). The "best professors" of scholarly and unscholarly students. Paper read at the annual meeting of the Western Psychological Association.
- Cobb, E. B. (1956). Construction of a forced-choice university instructor rating scale. Unpublished doctoral dissertation, University of Tennessee.
- Coffman, W. E. (1954). Determining students' concepts of effective teaching from their ratings of instructors. *Journal of Educational Psychology* 45:277-286.
- Corcoran, M. (1961). Where does college teaching stand in the career plans of superior college seniors? An exploratory study of interest in academic careers. Minneapolis, Minn.: Bureau of Institutional Research, University of Minnesota.
- Cosgrove, D. M. (1959). Diagnostic rating of teacher performance. *Journal of Educational Psychology* 50:200-204.
- Costin, F. (1971). Empirical test of the "teacher-centered" versus "student-centered" dichotomy. *Journal of Educational Psychology* 62:410-412.
- Costin, F. (1974). Measuring lecturing behavior of college instructors. *Professional Psychology* 5:106-108.
- Costin, F., Greenough, W. T., and Menges, R. J. (1971). Student ratings of college teaching: Reliability, validity, and usefulness. *Review of Educational Research* 41:511-535.
- Costin, F., and Grush, J. E. (1973). Personality correlates of teacher-student behavior in the college classroom. *Journal of Educational Psychology* 65:35-44.
- Crannell, C. W. (1948). An experiment in the rating of instructors by their students. *College and University* 24:5-11.
- Crannell, C. W. (1943). A preliminary attempt to identify the factors in student-instructor evaluation. *Journal of Psychology* 36:417-422.
- Crawford, P. L., and Bradshaw, H. L. (1968). Perception of characteristics of effective university teachers: A scaling analysis. *Educational and Psychological Measurement* 28:1079-1085.
- Creager, J. A. (1950). A multiple-factor analysis of the Purdue Rating Scale for Instructors. *Purdue University Studies in Higher Education* 70:75-96.
- Crittenden, K. S., and Norr, J. L. (1973). Student values and teacher evaluation: A problem in person perception. *Sociometry* 36:143-151.
- Crouch, H. B., and Leathshaw, C. M. (1951). The validity of student opinions in evaluating a program of college biology. *Science Education* 35:73-76.
- Dansereau, R. A. (1973). Some myths on student evaluation of faculty. *Improving College and University Teaching* 21:48.
- Davis, C. O. (1926). Our best teachers. *School Review* 34:754-759.
- Davis, R. H. (1969). Student Instructional Rating Systems (SIRS): Technical bulletin. East Lansing, Mich.: Michigan State University.
- Deegan, W. L. (1972). Should students evaluate faculty? *Community and Junior College Journal* 43:25-26.

- Deshpande, A. S., Webb, S. C., and Marks, E. (1970). Student perceptions of engineering instructor behaviors and their relationships to the evaluation of instructors and courses. *American Educational Research Journal* 7:289-305.
- Dick, W. (1967). *Course Attitude Questionnaire: Its development, uses, and research results*. Report no. 67-1 (revision of no. 106, as revised by D. Stickell). University Park, Penn.: Office of Examination Services, University Division of Instructional Services, Pennsylvania State University.
- Diener, T. J. (1973). Can the student voice help improve teaching? *Improving College and University Teaching* 21: 35-37.
- Doyle, K. O., Jr. (1972). Construction and evaluation of scales for rating college instructors. Unpublished doctoral dissertation, University of Minnesota.
- Doyle, K. O., Jr. (1975). "Student Evaluation of Instruction." Lexington, Mass.: Heath.
- Doyle, K. O., Jr., and Liu, H. C. (1972). Factor analysis of "Student Evaluation of Instruction, Part I." Minneapolis, Minn.: Measurement Services Center, University of Minnesota.
- Drayer, A. M. (1961). Students' views of the qualifications of their teachers. *Journal of Teacher Education* 12:338-341.
- Drucker, A. J., and Remmers, H. H. (1951). Do alumni and students differ in their attitudes toward instructors? *Journal of Educational Psychology* 42:129-143.
- Dwyer, F. (1968). A review of characteristics and relationships of selected criteria for evaluating teacher effectiveness. University Park, Penn.: University Division of Instructional Services. Pennsylvania State University.
- Eble, K. E. (1972). "Professors as Teachers." San Francisco: Jossey-Bass.
- Echandia, P. P. (1963). A methodological study and factor analytic validation of forced-choice performance ratings of college accounting instructors. Unpublished doctoral dissertation, New York University.
- Endeavor Information Systems, Inc. (1973). *Factor structure of the Endeavor Instructional Rating Form*. Chicago: Endeavor Information Systems, Inc.
- Feldman, K. A. (1971). Using the work of others: Some observations on reviewing and integrating. *Sociology of Education* 44:86-102.
- Feldman, K. A. (1976). Grades and college students' evaluations of their courses and teachers. *Research in Higher Education* 4:69-111.
- Fenker, R., and Secrest, L. (n.d.). Teacher evaluation at Texas Christian University: An analysis of the perceived roles of faculty by students, administrators, and faculty. Unpublished manuscript.
- Finkbeiner, C. T., Lathrop, J. S., and Schuerger, J. M. (1973). Course and instructor evaluation: Some dimensions of a questionnaire. *Journal of Educational Psychology* 64:159-163.
- French, G. M. (1957). College students' concept of effective teaching determined by an analysis of teacher ratings. Unpublished doctoral dissertation, University of Washington.
- French-Lazovik, G. (1974). Predictability of students' evaluations of college teachers from component ratings. *Journal of Educational Psychology* 66:373-385.
- Frey, P. W. (1973). Comparative judgment scaling of student course ratings. *American Educational Research Journal* 10:149-154.
- Frey, P. W. (1974). The ongoing debate: Student evaluation of teaching. *Change*, February pp. 47-48; 64.

- Fulcher, D. G., and Anderson, W. T., Jr. (1974). Interpersonal dissimilarity and teaching effectiveness: A relational analysis. *Journal of Educational Research* 68:19-25.
- Gadzella, B. M. (1967). Student view of an "ideal" professor. *College Student Survey* 1:7-8; 21.
- Gadzella, B. M. (1968a). College students' views and ratings of an ideal professor. *College and University* 44:89-95.
- Gadzella, B. M. (1968b). The ideal professor as viewed by New Paltz freshmen. Report No. 9. New Paltz, N.Y.: Office of Institutional Research, State University College, New Paltz.
- Gadzella, B. M. (1974). A bridge of understanding between students and faculty. *Improving College and University Teaching* 22:9-11.
- Gagné, F., and Allaire, D. (1974). Reliability and construct validity of a discrepancy approach to measuring student perceptions of teaching. Paper read at the annual meeting of the American Psychological Association.
- Garber, H. (1964). Certain factors underlying the relationship between course grades and student judgments of college teachers. Unpublished doctoral dissertation, University of Connecticut.
- Geyer, D. L. (1946). Qualities desired in college instructors. *School and Society* 63:270-271.
- Gibb, C. A. (1955). Classroom behavior of the college teacher. *Educational and Psychological Measurement* 15:254-263.
- Good, K. C. (1971). Similarity of student and instructor attitudes and student's attitudes toward instructors. Unpublished doctoral dissertation, Purdue University.
- Gottlieb, D. (1962). Social class, achievement, and the college-going experience. *School Review* 70:273-286.
- Graham, M. H. (1972). The relationship between CEQ ratings and instructor rank, class size, and course level. Research report no. 337. Measurement and Research Division, Office of Instructional Resources, University of Illinois.
- Granzin, K. L., and Painter, J. J. (1973). A new explanation for students' course evaluation tendencies. *American Educational Research Journal* 67:451-460.
- Grasha, A. F. (1975). The role of internal instructor frames of reference in the student rating process. *Journal of Educational Psychology* 67:451-460.
- Guilford, J. P. (1954). "Psychometric Methods," (2nd ed.). New York: McGraw-Hill.
- Gulo, E. V. (1971). Measuring dimensions of teaching effectiveness with the Semantic Differential. Paper read at the annual meeting of the Northeastern Educational Research Association.
- Gustad, J. W. (1967). Evaluation of teaching performance: Issues and possibilities. In "Improving College Teaching" (Lee, C. B. T., ed.). Washington, D.C.: American Council on Education.
- Haggard, W. W. (1943). Some freshmen describe the desirable college teacher. *School and Society* 58:238-240.
- Hall, D. T. (1970). The effect of teacher-student congruence upon student learning in college classes. *Journal of Educational Psychology* 61:205-213.
- Hanke, J. E. (1970). Teacher and student perceptions as predictors of college teacher effectiveness. Unpublished doctoral dissertation, University of Northern Colorado.

- Harman, H. H. (1967). "Modern Factor Analysis," 2nd ed. Chicago: University of Chicago Press.
- Harry, J., and Goldner, N. S. (1972). The null relationship between teaching and research. *Sociology of Education* 45:47-60.
- Hartley, E. L., and Hogan, T. P. (1972). Some additional factors in student evaluation of courses. *American Educational Research Journal* 9:241-250.
- Hartung, A. B. (1972). Teaching excellence. *Improving College and University Teaching* 20:146-147.
- Harvey, J. N., and Barker, D. G. (1970). Student evaluation of teaching effectiveness. *Improving College and University Teaching* 18:275-278.
- Hayes, R. B. (1963). A way to measure classroom teaching effectiveness. *Journal of Teacher Education* 14:168-176.
- Heidgerken, L. E. (1952). "The Nursing Student Evaluates Her Teachers." Philadelphia, Penn.: Lippincott.
- Heyn, D. R. (1972). **Development and Validation of a Sociometric Instructor Evaluation Instrument and Procedure.** U. S. Department of Health, Education, and Welfare, Office of Education, National Center for Educational Research and Development, Final Report. Sherman, Texas: Austin College. (ERIC document reproduction service no. ED 068 077).
- Hildebrand, J., Wilson, R. C., and Dienst, E. R. (1971). *Evaluating University Teaching.* Berkeley, Calif.: Center for Research and Development in Higher Education, University of California, Berkeley.
- Hill, A. M. (1945). How do *you* do? *Journal of Engineering Education* 36:209-210.
- Hodgson, T. F. (1958). The general and primary factors in student evaluation of teaching. Unpublished doctoral dissertation, University of Washington.
- Hoffman, R. W. (1963). Students portray the excellent teacher. *Improving College and University Teaching* 11:21-24.
- Holmes, D. S. (1971). The Teaching Assessment Blank: A form for the student assessment of college instructors. *Journal of Experimental Education* 39:34-38.
- Hoyt, D. P. (1973). Identifying effective educational procedures. *Improving College and University Teaching* 21:73-76.
- Hussain, K. M., and Leestamper, R. (1968). **Survey on Criteria of Teaching Effectiveness at New Mexico State University.** Las Cruces, N.M.: New Mexico State University. (ERIC document reproduction service no. ED 023 365).
- Isacson, R. L., McKeachie, W. J., Milholland, J. E., Lin, Y. G., Hofeller, M., Baerwaldt, J. W., and Zinn, K. L. (1964). Dimensions of student evaluations of teaching. *Journal of Educational Psychology* 55:344-351.
- Jaeger, R. M., and Freijo, T. D. (1974). Some psychometric questions in the evaluation of professors. *Journal of Educational Psychology* 66:416-423.
- Jenkins, C., Baker, R., Emerson, D., Hagerty, D., and Tune, D. (1970). *Evaluation of community college instruction: A background study.* Freeport, Ill.: Highland Community College.
- Jiobu, R. M., and Pollis, C. A. (1971). Student evaluations of courses and instructors. *American Sociologist* 6:317-321.
- Kennedy, W. R. (1975). Grades expected and grades received — their relationship to students' evaluations of faculty performance. *Journal of Educational Psychology* 67:109-115.
- Kinnane, M. (1961). *Attitudes of Undergraduates Toward College Teaching as a career.* Winchester, Mass.: New England Board of Higher Education.

- Kohlan, R. G. (1973). A comparison of faculty evaluations early and late in the course. *Journal of Higher Education* 44:587-595.
- Kreuger, L. M. (1936). Traits in college teachers preferred by students: A study in leadership. *Phi Delta Kappan* 19:14-19.
- Krupka, J. G. (1970). Report on Faculty and Student Evaluation of Instructor Rating Questionnaire. Northampton, Penn.: Northampton County Area Community College (ERIC document reproduction service no. ED 038 964).
- Kulik, J. A., and McKeachie, W. J. (1975). The evaluation of teachers in higher education. In "Review of Research in Education" (Kerlinger, F. N., ed.), vol. 3.
- Lahat-Mandelbaum, B., and Kipnis, D. (1973). Leader behavior dimensions related to students' evaluation of teaching effectiveness. *Journal of Applied Psychology* 58:250-253.
- Lamson, E. E. (1942). Some college students describe the desirable college teacher. *School and Society* 56:615.
- Lathrop, R. G. (1968). Unit factorial ratings by college students of courses and instructors. Paper read at the annual meeting of the Western Psychological Association.
- Leftwich, W. H., and Remmers, H. H. (1962). A comparison of graphic and forced-choice ratings of teaching performance at the college and university level. *Purdue University Studies in Higher Education* 92:3-31.
- Lehmann, I. J. (1966). The college faculty as perceived by graduates and non-graduates. *Journal of Educational Measurement* 3:169-173.
- Leonard, W. M., II. (1973). Student preferences for what makes a good college teacher. *Improving College and University Teaching* 21:10-13.
- Levinthal, C. F. (1974). An analysis of the teacher evaluation process. Final Report, U.S. Department of Health, Education, and Welfare, National Institutes of Education, project no. 2B089. Hempstead, N.Y.: Hofstra University.
- Levinthal, C. F., Lansky, L. M., and Andrews, O. E. (1971). Student evaluations of teacher behaviors as estimations of real-ideal discrepancies. A critique of teacher rating methods. *Journal of Educational Psychology* 62:104-109.
- Ludeman, W. W. (1960). A study of the qualities of the ideal college teacher. *Teachers College Journal* 31:125-126.
- Mass, J. B., and Owen, T. R. (1973). Cornell Inventory for Student Appraisal of Teaching and Courses: Manual of Instructions. Ithaca, N.Y.: Center for Improvement of Undergraduate Education, Cornell University.
- MacDonald, M. E. (1931). Students' opinions as regards desirable and undesirable qualifications and practices of their teachers in teacher-training institutions. *Educational Administration and Supervision* 17:139-146.
- Mann, R. D., Arnold, S. M., Binder, J. L., Cytrynbaum, S., Newman, B. M., Ringwald, B. E., Ringwald, J. W., and Rosenwein, R. (1970). "The College Classroom: Conflict, Change, and Learning." New York: Wiley.
- Mann, W. R. (1968). Changes in the level of attitude sophistication of college students as a measure of teacher effectiveness. Unpublished doctoral dissertation, University of Michigan.
- Marshall, M. S. (1952). Grading the teacher. *Association of American Colleges Bulletin* 38:257-267.
- McComas, J. D. (1965). Profiles of teachers. *Improving College and University Teaching*, 13:135-136.
- McKeachie, W. J. (1972). Student evaluations keyed to function. Paper read at the annual meeting of the American Educational Research Association.

- Meinkoth, M. R. (1971). Teachers of economic principles: Effect on student achievement and attitudes. *Journal of Experimental Education* 40:66-72.
- Meredith, G. M. (1969). Dimensions of faculty-course evaluation. *Journal of Psychology* 73:27-32.
- Miklich, D. R. (1969). An experimental validation study of the Purdie Rating Scale for Instruction. *Educational and Psychological Measurement* 29:963-967.
- Miller, R. I. (1972). "Evaluating Faculty Performance." San Francisco: Jossey-Bass.
- Miller, R. I. (1974). "Developing Programs for Faculty Evaluation." San Francisco: Jossey-Bass.
- Morton, R. D. (1965). Students views of teaching. *Improving College and University Teaching* 13:140-142.
- Mueller, R. H., and Roach, P. J. (1970). What constitutes an "ideal professor"? Unpublished manuscript.
- Mueller, R. H., Roach, P. J., and Malone, J. A. (1971). College students' views of the characteristics of an "ideal" professor. *Psychology in the Schools* 8:161-167.
- Musella, D., and Rusch, R. (1968). Student opinion on college teaching. *Improving College and University Teaching* 16:137-140.
- Naftulin, D. H., Ware, J. E., Jr., and Donnelly, F. A. (1973). The Doctor Fox lecture: A paradigm of educational seduction. *Journal of Medical Evaluation* 48:630-635.
- Neeley, M. (1968). A teacher's view of teacher evaluation. *Improving College and University Teaching* 16:207-209.
- Newcomb, T. M., Turner, R. H., and Converse, P. E. (1965). "Social Psychology: The Study of Human Interaction." New York: Holt, Rinehart, and Winston.
- Odom, C. L. (1943). An objective determination of the qualities of a good college teacher. *Peabody Journal of Education* 21:109-116.
- Owen, P. H. (1967). Some dimensions of college teaching: An exploratory study using critical incidents and factor analyses of student ratings. Unpublished doctoral dissertation, University of Houston.
- Perlman, D. (1973). Class size and students' ratings of university courses. Paper read at the annual meeting of the Canadian Psychological Association.
- Permut, S. E. (1973). Cue utilization patterns in student-faculty evaluation. *Journal of Psychology* 83:41-48.
- Perry, R. R. (1969). Evaluation of teaching behavior seeks to measure effectiveness. *College and University Business* 47:18-22.
- Perry, R. R., and Baumann, R. R. (1973). Criteria for the evaluation of college teaching: Their reliability and validity at the University of Toledo. In (Sockloff, A. L., ed.). *Proceedings of the First Invitational Conference on Faculty Effectiveness as Evaluated by Students*. Philadelphia, Penn.: Measurement and Research Center, Temple University.
- Pipes, W. H. (1951). College students "flunk" teachers. *Association of American Colleges Bulletin* 37:263-265.
- Plant, W. T., and Sawrey, J. M. (1970). Student ratings of psychology professors as teachers and the research involvement of the professors rated. *Clinical Psychologist* 23:15-16; 19.
- Pogue, F. G., Jr. (1967). Students' ratings of the "ideal teacher." *Improving College and University Teaching* 15:133-136.

- Pohlmann, J. T. (1975). A description of teaching effectiveness as measured by student ratings. *Journal of Educational Measurement* 12:49-54.
- Price, J. A., and Magoon, A. J. (1971). Predictors of college student ratings of instructors. Paper read at the annual meeting of the American Psychological Association.
- Purohit, A., and Magoon, A. J. (1971). The validity of student-run course evaluations. Paper read at the annual meeting of the American Educational Research Association.
- Quereshi, M. Y., and Widlak, F. W. (1973). Students' perception of a college teacher as a function of their sex and achievement level. *Journal of Experimental Education* 41:53-57.
- Quick, A. F., and Wolfe, A. D. (1965). The ideal professor. *Improving College and University Teaching* 13:133-134.
- Remmers, H. H. (1929). The college professor as the student sees him. *Purdue University Studies in Higher Education* 11:1-63.
- Remmers, H. H., and Weisbrodt, J. A. (1964). "Manual of Instructions for Purdue Rating Scale of Instruction." Purdue, Ind.: Purdue Research Foundation.
- Riley, J. W., Ryan, B. F., and Lifshitz, M. (1950). "The Student Looks at His Teacher: An Inquiry Into the Implications of Student Ratings at the College Level." New Brunswick, N.J.: Rutgers University Press.
- Riley, J. W., Ryan, B. F., and Lifshitz, M. (1950). *The Student Looks at His Teacher: An Inquiry Into the Implications of Student Ratings at the College Level.* New Brunswick, N.J.: Rutgers University Press.
- Romine, S. (1974). Student and faculty perceptions of an effective university instructional climate. *Journal of Educational Research* 68:139-143.
- Rosenshine, B., Cohen, A., and Furst, N. (1973). Correlates of student preference ratings. *Journal of College Student Personnel* 14:269-272.
- Sagen, H. B. (1974). Student, faculty, and department chairmen ratings of instructors: Who agrees with whom? *Research in Higher Education* 2:265-272.
- Schubert, D. G. (1953). Student criticisms of college teaching. *Journal of Teacher Education* 4:294.
- Shane, H. G. (1965). How do they rate you professor? *NEA Journal* 54:18-20.
- Sherman, B., and Blackburn, R. T. (1975). Personal characteristics and teaching effectiveness of college faculty. *Journal of Educational Psychology* 67:124-131.
- Singhal, S., and Stallings, W. M. (n.d.). A study of the relationships between course evaluations by students and severity of grading by instructors in freshman rhetoric at the University of Illinois. Research report no. 252, Urbana, Ill.: Measurement and Research Division, Office of Instructional Resources, University of Illinois.
- Smith, A. A. (1944). What is good college teaching? *Journal of Higher Education* 15:216-218.
- Smith, A. A. (1948). "He who can, teaches." *School and Society* 68:188-189.
- Sockloff, A. L., and Deabler, V. T. (1971). The construction of the Faculty and Course Evaluation Instrument. Research report 71-2. Philadelphia, Penn.: Testing Bureau, Temple University.
- Solomon, D. (1966). Teacher behavior dimensions, course characteristics, and student evaluations of teachers. *American Educational Research Journal* 3:35-47.
- Spaights, E. (1967). Students appraise teachers' methods and attitudes. *Improving College and University Teaching* 15:15-17.
- Spencer, R. E. (1967). Analysis of the Instructor Rating Form - General Engineering Department. Research report no. 253. Urbana, Ill.: Measurement and Research Division, Office of Instructional Resources, University of Illinois.

- Spencer, R. E. (1969). Some dimensions of the Illinois Course Evaluation Questionnaire. Research report no. 303. Urbana, Ill.: Measurement and Research Division, Office of Instructional Resource, University of Illinois.
- Stalnaker, J. M., and Remmers, H. H. (1928). Can students discriminate traits associated with success in teaching? *Journal of Applied Psychology* 12:602-610.
- Sullivan, A. M., and Skanes, G. R. (1974). Validity of student evaluation of teaching and the characteristics of successful instructors. *Journal of Educational Psychology* 66:584-590.
- Taguiri, R. (1969). Person perception. In "The Handbook of Social Psychology" (Lindzey, G., and Aronson, E., eds.). 2nd ed., volume 3: "The Individual in a Social Context." Reading Mass.: Addison-Wesley.
- Taylor, G. F. (1959). Characteristics of best liked teacher, least liked teacher, and most effective teacher in teacher training institutions. Unpublished doctoral dissertation, University of Connecticut.
- Taylor, R. E. (1968). An investigation of the relationship between psychological types in the college classroom and the student perception of the teacher and preferred teaching practices. Unpublished doctoral dissertation, University of Maryland.
- Thorndike, R. L., and Hagen, E. (1969). "Measurement and Evaluation in Psychology and Education," 3rd ed. New York: Wiley.
- Tobias, S., and Hanlon, R. E. (1974). Attitudes Towards Instructors, Social Desirability, and Behavioral Intentions. New York: Division of Teacher Education, City University of New York. (ERIC document reproduction service no. ED 101 016).
- Tobias, S., and Hanlon, R. (1975). Attitudes towards instructors, social desirability, and behavioral intentions. *Journal of Educational Psychology* 67:405-408.
- Turner, R. L. (1970). Good teaching and its contexts. *Phi Delta Kappan* 52:155-158.
- Turner, R. L., Evans, J. H., Hale, T. A., Cairns, S. G., and Maleski, F. D. (1969). How do student characteristics affect their evaluations of instructors. *Bulletin of the School of Education, Indiana University* 45:47-97.
- Van Horn, C. (1968). An analysis of the 1968 course and instructor evaluation report. Institutional research bulletin no. 2-68. West Lafayette, Ind.: Measurement and Research Center, Purdue University.
- Veldman, D. J. (1968). Student evaluation of college of education courses, fall semester, Unpublished manuscript.
- Villano, M. W. (1975). The relationship of certain course characteristics to student ratings of science and mathematics teaching at four-year and two-year colleges. Paper read at the annual meeting of the American Educational Research Association.
- Villano, M. W., and Rosenstock, E. H. (1973). A decade with the Course Attitude Questionnaire: A factorial study. Paper read at the annual meeting of the Northeastern Educational Research Association.
- Voecks, V. (1954-55). Ridicule and other detriments to effective teaching. *American Association of University Professors Bulletin* 40:621-630.
- Walker, B. D. (1968). An investigation of selected variables relative to the manner in which a population of junior college students evaluate their teachers. Unpublished doctoral dissertation, University of Houston.
- Walsh, G. V. (1972). One in five made us think. *Improving College and University Teaching* 20:153-155.

- Whitely, S. E., Doyle, L. O., Jr., and Hopkinson, K. (1973). Student ratings and criteria for effective teaching. Report 731 F. Minneapolis, Minn.: Measurement Services Center, University of Minnesota.
- Whitlock, G. H. (1971). Evaluating instruction: Learning/perception. *Teaching-Learning Issues*, no. 16. Knoxville, Tenn.: Learning Research Center, University of Tennessee.
- Widlak, F. W., McDaniel, E. D., and Feldhusen, J. F. (1973). Factor analysis of an instructor rating scale. Paper read at the annual meeting of the American Educational Research Association.
- Widlak, F. W., and Quereshi, M. Y. (1972). Student characteristics and instructor ratings: A person-perception approach. Paper read at the annual meeting of the American Psychological Association.
- Williams, H. Y., Jr. (1965). College students' perceptions of the personal traits and instructional procedures of good and poor teachers. Unpublished doctoral dissertation, University of Minnesota.
- Withington, R. (1944) On judgment and grades. *American Association of University Professors Bulletin* 30:557-568.
- Wortruba, T. R., and Wright, P. L. (1975). How to develop a teacher-rating instrument: A research approach. *Journal of Higher Education* 46:653-663.
- Yonge, G. D., and Sassenrath, J. M. (1968). Student personality correlates of teacher ratings. *Journal of Educational Psychology* 59:44-52.
- Yourglic, A. (1955). Study on correlations between college teachers' and students' concepts of "ideal-student" and "ideal-teacher." *Journal of Educational Research* 49:59-64.

APPENDIX A

Brief descriptions of 49 studies, data from which form the basis of Table I, are given below. Studies are numbered from S1 to S49 (with separate numbers, when appropriate, for each of two different studies given in the same article or report). At the end of each description are four numbers in parenthesis. The first of these (preceding the colon) is the number of characteristics in the study that were ranked for the present analysis; following the colon is the number of characteristics coded into two (or, in one instance, in Hoffman, 1963, more than two) of the coding categories of this analysis, the number of characteristics coded as "other professional characteristics," and the number of characteristics coded as "personal characteristics" (see footnote 1), respectively.

S1 Birney, Coplin and Grose (1960): seniors at Amherst College describing their best teacher in each division of three division—Humanities, Social Sciences, Natural Sciences (Spring, 1959). The ranking given in the report (Table 26) is based on the overall frequency of mention of each characteristic. (9:4,0,1)

S2 Bousfield (1940): 61 students listing five traits or qualities they regard as most desirable in college professors. The ranking given in the article (Table 1) is based on the frequency with which each trait or quality was mentioned. (18:0,1,5)

S3 Bousfield (1940): 310 students at Tufts and 197 students at the University of Connecticut rating 19 qualities as to their desirability or importance in a professor.

The ranking given in the article (Table 2) is based on the average rating of each quality. (19:1,2,4)

S4 Brewer and Brewer (1970): students from De Paul University (N = 392), from Northwestern University (N = 137), and from North Park College (N = 118) rating 10 qualities as to their importance for good college teaching. The ranking in the present analysis is based on the "scale values" of each characteristic, as given in Table 3 in the article. (10:1,1,2)

S5 Brewer and Brewer (1970): 120 students at De Paul University rating 10 qualities as to the degree to which they were characteristic of some specific, excellent teacher. The ranking in the present analysis is based on the mean rating of each characteristic, as given in Table 4 in the article. (10:1,1,2)

S6 Bridges, Ward, Brown and Greenwood (1971): 502 students at "a large university" listing six outstanding characteristics of the best college teachers they have known. The ranking in the present analysis is based on the frequency with which responses fell into the same category of characteristics, as given in Table 2 in the article. (24:7,3,5)

S7 Case (1952): 209 engineering students divided among six universities describing the characteristics they thought made an outstanding teacher. The ranking in the present analysis is based on the percentage of students listing the same characteristic, as given in Table 4 of the article. Of the 32 characteristics listed, only the highest ranked 25 are included in the present analysis (25:2,2,6)

S8 Clinton (1930): 177 juniors listing the qualities they desired in their college teachers. The ranking in the present analysis is based on the frequency with which students mentioned a quality, as given on p. 702 in the article. Of the 35 qualities listed, only the highest ranked 25 are included in the present analysis. (25:0,4,9)

S9 Crawford and Bradshaw (1968): 158 students in four psychology classes "at a university" rating characteristics in terms of the degree to which they are essential or critical to effective university teaching. The ranking given in the article (Table 1) is based on the median scale value of each characteristic. (13:2,1,2)

S10 Drayer (1961): 148 students at a "four-year liberal arts college for men" stating their reasons for picking certain teachers as those they liked best. The ranking in the present analysis is based on the frequency with which responses fell into the same category of reasons, as given in Table 1 in the article. (16:1,2,3)

S11 Drucker and Remmers (1951): 251 students at Purdue University ranking the 10 traits of the Purdue Rating Scale for Instructors as to their importance to good college teaching (during early 1949). The ranking given in the article (Table 3) is based on the median ranking of each trait. (10:2,0,4)

S12 Gadzella (1968a): 443 students at Western Washington State College ranking the five most important criteria of an ideal professor from a list of 25 criteria (Spring, 1966). The ranking in the present analysis is based on the weighted ratings of the criteria, as given in Table 2 of the article (cf. Gadzella, 1967). (25:2,5,2)

S13 Gadzella (1968b): 1,145 entering freshmen at New York State College at New Paltz ranking the five most important criteria of an ideal professor from a list of 25 criteria (during Orientation classes, Fall, 1968). The ranking in the present analysis is based on the weighted ratings of the criteria, as given in Table 1 in the article. (25:2,5,2)

S14 Gadzella (1974): "three groups of students totalling 300" at Western Washington State College ranking the five most important criteria of an ideal professor (from a list of 17 criteria). The ranking in the present analysis is based on the weighted ratings of the criteria, as given in Table 1 in the article. (17:2,3,2)

S15 Geyer (1946): students at Chicago Teachers College ranking eight qualities as to their desirability in college instructors. The ranking is given in Table 1 in the article. (8:1,0,3)

S16 Gottlieb (1962): sample of entering freshmen males at a "large midwestern state university" (N = 238) and at a "smaller private institution" (N = 115) selecting from 15 characteristics the most important characteristic of a good college teacher. The general ranking in the present analysis has been calculated from the data given in Table 3 of the article. (15:0,8,0)

S17 Haggard (1943): 49 freshmen in the Western Washington College of Education (Bellingham) listing eight qualities they most desired in college teachers (during the 1943 winter quarter). The ranking in the present analysis is based on the frequency with which students mentioned the same quality, as given in Table 1 in the article. (22:0,5,6)

S18 Haggard (1943): 49 freshmen in the Western Washington College of Education (Bellingham) ranking six of eight qualities as to their desirability in college teachers (during the 1943 winter quarter). The ranking in the present analysis is based on the frequency with which a characteristic fell into the first six ranks, as given in Table 2 in the article. (8:1,0,3)

S19 Hayes (1963): 1,070 students at The Pennsylvania State University rating their best instructor on 14 attributes (during 1960). The ranking in the present analysis is based on the percentage of students endorsing a positive attribute and one minus the percentage endorsing a negative attribute (see Table 1 in the article). (14:1,4,0)

S20 Heidgerken (1952): 384 senior nursing students in 37 schools of nursing located in 21 states stating their reasons for their choice of the best teacher they have had in training (Spring, 1951). The ranking in the present analysis is based on the percentage of responses falling into the same category of reasons, as given in Tables 10-14 and on pp. 65, 72, 75, and 76. Of the 33 categories of responses, only the highest ranked 25 are included in the present analysis. (25:5,2,2)

S21 Heyn (1972): 254 students at Austin College stating their reasons for picking certain professors as one of the three professors they would most likely recommend to other students (Spring, 1972). The ranking in the present analysis is based on the number of students stating the same reason, as given in Appendix B, pp 58-59, in the report. Of the 49 reasons listed, only the highest ranked 25 are included in the present analysis. (25:3,3,2)

S22 Hoffman (1963): seniors at Hofstra College stating their reasons for picking certain teachers as outstanding. The ranking given in the article is based on the percentage of students giving reasons falling into the same category. (8:4,2,1)

S23 Hussein and Leestamper (1968): 283 students at New Mexico State University rating 60 "behavior criteria" as to their importance for teaching effectiveness (Fall, 1967). The ranking given in the report (Tables 1 and 2 in the report, taken together) is based on the weighted frequency for each criterion. Of the 60 criteria, only the highest ranked 25 are included in the present analysis. (25:1,5,0)

S24 Jenkins, Baker, Emerson, Hagerty and Tune (1970): students at Highland Community College (Freeport, Ill.) rating 60 "criterion statements" as to their importance to good instruction. The ranking given in the report (on pp. 20-22) is based on the average rating of each criterion statement. Of the 60 statements, only the highest ranked 25 are included in the present analysis. (25:1,6,1)

S25 Kinnane (1961): 3,418 students (with B or better averages) at "45 colleges and universities in six New England states" selecting the one most outstanding characteristic, from a list of five characteristics, of the best college teacher they have known. The ranking in the present analysis is based on the percentage of students selecting a particular characteristic, as given in Table 7 in the article. (5:1,1,0)

S26 Kreuger (1936): 1,085 students at 10 Indiana colleges and universities rating 60 traits as to their importance in a college teacher. The ranking of these traits is given on p. 17 of the article. Of the 60 traits, only the highest ranked 25 are included in the present analysis. (25:0,1,9)

S27 Krupka (1970): 60 students at Northampton County Area Community College ranking 10 specific "rating areas" as to their importance in the judgment of a teacher. The ranking given in the report (in the table on p. 2) is based on the average rating of each area. (Note: two global areas, "general estimate of teacher," and "general estimate of the course," are not included in the present analysis.) (10:0,1,1)

S28 Lamson (1942): "a group of students who were entering their senior year at the New Jersey State Teachers College, Jersey City" ranking six of eight qualities as to the degree to which students valued them in their teachers (Fall, 1938). The ranking in the present analysis is based on the frequency with which a quality fell into the first six ranks, as given in Table 1 in the article. (8:1,0,3)

S29 Lehmann (1966): 1,777 seniors at Michigan State University rating 15 qualities as to whether or not they are characteristic of a good college teacher (Spring, 1962). The general ranking in the present analysis has been calculated from the data given in Table 1 in the article. (15:0,9,0)

S30 Ludeman (1960): over 500 students at Southern State Teachers College (Springfield, South Dakota) ranking 10 traits of the ideal teacher (during the spring quarter of 1959). The ranking is given in the Table on p. 125 of the article (10:1,2,3)

S31 Mueller and Roach (1970): 225 students enrolled in psychology courses at the University of Windsor ranking the five most important criteria of an ideal professor

from a list of 25 criteria (during the 1969 summer session). The ranking in the present analysis is based on the weighted ratings of the criteria, as given in Table 1 of the article. (25:2,5,2)

S32 Mueller, Roach and Malone (1971): 642 introductory psychology students at the University of Windsor ranking the five most important criteria of an ideal professor from a list of 25 criteria (Fall, 1969). The ranking in the present analysis is based on the weighted ratings of the criteria, as given in Table 1 of the article. (25:2,5,2)

S33 Musella and Rusch (1968): 394 seniors at the State University of New York at Albany describing teacher behaviors that promoted their thinking (Spring, 1966). The ranking in the present analysis is based on the frequency with which a particular category of behavior was mentioned, as given in Table 2 of the article. (Note: two categories, "miscellaneous," and "blank," are not included in the present analysis.) (15:2,1,2)

S34 Musella and Rusch (1968): 394 seniors at the State University of New York at Albany ranking 10 qualities as to their importance for teaching in general (Spring, 1966). The ranking is based on the frequency with which each quality was chosen, as given in Table 4 in the article. (10:0,0,1)

S35 Owen (1967): 42 undergraduate students in a sophomore level political science class at the University of Houston describing some specific thing that a very effective teacher did that made for particularly successful instruction; 78 undergraduate students in three separate classes in industrial psychology (University of Houston) describing some specific thing that an instructor in an outstandingly good course did that was important to making the course superior; and 15 undergraduate students in a sophomore level accounting class (University of Houston) describing some specific thing an instructor in an outstandingly good, large class, and that an instructor in an outstandingly good, small class, did to make the particular course superior. The ranking in the present analysis is based on the number of descriptions falling into the same category (excluding various miscellaneous categories), as given in Table 1 (Col. 11) in the report. Of the 35 usable categories, only the highest ranked 23 are included in the present analysis (there was a 5-way tie for rank 24). (23:1,0,1)

S36 Permut (1973): 14 students in a "summer session survey-type course at the University of Illinois" ranking the relative importance of 10 traits in arriving at their overall effectiveness evaluations of hypothetical "instructor profiles." The ranking given in the article (Table 1) is based on the average rank of each trait. (10:0,0,0)

S37 Perry (1969): "a sample of students stratified by college and class rank" at the University of Toledo rating 60 "criterion statements" as to their importance for the evaluation of effective teaching behavior. The ranking given in the article (in the Table on pp. 18-19) is based on the weighted raw score total for each criterion. Of the 60 statements, only the highest ranked 25 are included in the present analysis. (25:1,6,0)

S38 Pogue (1967): 307 students (of 530 students contacted) at Philander Smith College ranking 10 statements describing the ideal professor. Students were asked

to discriminate among statements as to the first, second and third in importance. The ranking given in the article (Table 2) is based on the weighted rating of each statement. (10:0,1,0)

S39 Quick and Wolfe (1965): 483 students at the University of Oregon ranking 10 statements describing the ideal professor (during the winter quarter of 1963). Students were asked to discriminate among the statements as to the first, second and third in importance. The ranking given in the article (Table 1) is based on the weighted rating of each statement. (10:0,0,2)

S40 Riley, Ryan and Lifshitz (1950): 6,681 students at Brooklyn College selecting from 10 qualities the three most important to good teaching in the arts, the sciences, and the social sciences (near the end of the academic year 1946-1947). The general ranking in the present analysis has been calculated from data given in Table 4 in the report. (10:0,0,2)

S41 Schubert (1953): 100 students at Los Angeles State College of Applied Arts and Sciences listing the traits and qualities they liked in college teachers. The ranking given in the article (pp. 97-98) is based on the frequency with which a trait or quality was mentioned. (8:0,0,2)

S42 Smith (1944): 100 students describing their ideal teacher. The ranking given in the article (pp. 216-217) is based on the frequency of mention of each trait description. Of the 26 categories of traits, only the highest ranked 25 are included in the present analysis. (25:6,4,10)

S43 Smith (1948): 100 senior and graduate students in a course in educational sociology at Purdue University ranking 10 traits of the ideal university teacher. The ranking is given in Table 1 in the article. (10:0,0,2)

S44 Stalnaker and Remmers (1928): 26 students in a class at Purdue University ranking the ten traits of the Purdue Rating Scale for Instructors as to their importance in the teaching situation. The ranking in the present analysis is based on summing the sums of ratings for each of the two random halves of the class, as given in Table 1 of the article. (Note: the resultant ranking is given directly in Remmers, 1929, p 19.) (10:2,0,4)

S45 Taylor (1959): 351 seniors and 444 graduate students at 21 colleges and universities describing their best-liked teacher. The ranking in the present analysis is based on the frequency with which responses fell in the same category (across two coders), as given in Table 14 in the report. (20:2,1,3)

S46 Taylor (1959): 351 seniors at 15 colleges and universities selecting any of those characteristics possessed by their best liked teacher (from a list of 60 characteristics). The ranking in the present analysis is based on the frequency with which male and female students selected a given characteristic, as given in Table 19 in the report. Of the 60 characteristics, only the highest ranked 25 are included in the present analysis. (25:1,4,5)

S47 Walsh (1972): 117 students at "State University College" and 184 students at

“State University Center” selecting three qualities of which they were most aware in the instructor having the greatest impact on their thinking processes (from a list of 13 qualities). The ranking in the present analysis is based on averaging the rankings given separately for students at the two different schools, as given in Table 2 in the article. (13:0,1,0)

S48 Williams (1965): 777 juniors and seniors at three “private liberal arts colleges” in the Twin Cities area (Minnesota) describing the single personal characteristic of the best teacher that stood out most in their minds (Spring, 1963). The ranking in the report (pp. 79-80) is based on the frequency with which a given characteristic was used. (8:0,0,1)

S49 Younglich (1955): a sample of 101 students at a “university in the Pacific Northwest” listing traits which they thought comprised the “ideal teacher.” The ranking in the present analysis is based on the frequency with which responses fell into the same category, as given in Table 2 in the article. (18:0,2,9)

Each of the following studies contains information about instructor traits that students consider to be important to good teaching or characteristic of their ideal or best teachers, but none of these studies is included as part of Table I for one or another reason (students' responses about primary-school and secondary-school teachers are not separated out from their responses about college teachers; students' responses are combined with those of faculty or administrative personnel; less than five characteristics are given in the study; students have been asked to indicate their preferences on multi-item scales rather than to respond to single items; the information given in the study is incomplete or the data are presented in such a way that the characteristics cannot be ranked): Asher (1970), Blai (1974), Bogardus (1946), Breed (1927), Champlin (1928, 1953), Cloer (1970), Costin and Grush (1973), Crouch and Leathers (1951), Davis (1926), Echandia (1973), Fenker and Secrest (n.d.), Gulo (1971), Hall (1970), Hartung (1972), Hill (1945), Leonard (1973), MacDonald (1931), McComas (1965), Morton (1965), Odom (1943), Pipes (1951), Shane (1965), Spaight (1967), Taylor (1968), Turner, Evans, Hale, Cairns, and Maleski (1969; cf. Turner, 1970), and Whitlock (1971). Studies by Grasha (1975), Romine (1974), and Wortruba and Wright (1975) were located too late to be included as part of Table I.

APPENDIX B

Brief descriptions of 23 studies, data from which form the basis of Table II, are given below. Studies are numbered from S50 to S72. At the end of each description are three numbers in parenthesis: the number of specific evaluation items coded into two of the coding categories of the present analysis, the number of characteristics coded as “other professional characteristics,” and the number of characteristics coded as “personal characteristics” (see footnote 1), respectively.

S50 Brooks, Tarver, Kelley, Liberty and Dickerson (1971): approximately 10,000 rating forms completed by students in 318 classes taught by 163 different instructors in the College of Arts and Sciences at the University of Texas at Austin (Spring, 1968). The ranking in the present analysis is based on the size of the correlation between the overall rating of the instructor (“comparison of instructor with other

instructors'') and each of 25 specific evaluation items, as given in Table 2 in the report. (1,2,0)

S51 Centra (1975): student ratings of 78 teachers at a "new, small college." The ranking in the present analysis is based on the size of the correlation between the overall rating of the instructor ("instructor's effectiveness in course") and each of 15 specific evaluation items, as given in Table 2 in the article. (0,3,0)

S52 Cobb (1956): about 600 students in 124 sophomore, junior, or senior classes at 70 different colleges and universities, each rating a college teacher of his or her choice. The ranking in the present analysis is based on the size of the correlation between the overall rating of the instructor ("effectiveness as a college teacher") and each of 25 specific evaluation items, as given in Appendix F, data for Form M, in the report. Only the 25 largest of the 66 correlations have been ranked. (1,5,3)

S53 French-Lazovik (1974; cf. French, 1957): 3,654 student ratings of 133 faculty at the University of Washington (1956-57 academic year). The ranking in the present analysis is based on the size of the correlation of the overall rating of the instructor ("effectiveness as a college teacher") and each of 25 specific evaluation items, as given in Table 1 in the article. Only the 25 largest of the 41 correlations have been ranked. (3,2,1)

S54 French-Lazovik (1974): 6,120 student ratings of 144 teachers at the University of Pittsburgh (1971-72 academic year). The ranking in the present analysis is based on the size of the correlation of the overall rating of the instructor ("effectiveness as a college teacher") and each of 16 specific evaluation items, as given in Table 2 in the article. (0,1,0)

S55 Garber (1964): 430 students rating as many teachers at the University of Connecticut (June, 1963). The ranking in the present analysis is based on the size of the correlation of the overall rating of the instructor ("overall summary as a teacher") and each of 7 specific evaluation items, as given in Appendix B of the report (see the "student-as-unit" columns). (1,0,2)

S56 Good (1971): 409 students rating 14 instructors teaching 21 different sections of an undergraduate educational psychology course at Purdue University (main campus and four branch campuses). The ranking in the present analysis is based on the size of the correlation of the overall rating of the instructor ("desire to have instructor again") and each of 12 specific evaluation items, as given in Table 1 in the report. (0,0,4)

S57 Harry and Goldner (1972): student ratings of teachers in the College of Liberal Arts and Education at a "large, public urban midwestern university" (spring quarter, 1968-69). The ranking in the present analysis is based on the size of the correlation between the overall rating of the instructor (percent giving the instructor an "A" rating) and each of 5 specific evaluation items, as given in Table 2 in the article. (0,0,0)

S58 Harvey and Barker (1970): 37 undergraduate and 22 graduate male students living in college-operated apartments for married students rating certain of their

teachers (during the semester in which the rating was done). The ranking in the present analysis is based on the size of the point bi-serial correlation between the overall rating of the instructor (classification as either the "most effective" or "least effective" teacher) and each of 19 specific evaluation items, as given in Table 1 in the article. (1,2,2)

S59 Jioubu and Pollis (1971): student ratings of 67 courses at the University of Wisconsin, Green Bay (Fall, 1969). The ranking in the present analysis is based on the size of the correlation between the overall rating of the course ("how good a course") and each of 12 specific evaluation items, as given in the table on p. 320 in the article. (1,1,0)

S60 Leftwich and Remmers (1962): 2,109 student ratings of 80 instructors in 111 classes at a "large midwestern university." The ranking in the present analysis is based on the correlation between the total rating score for the instructor (across 10 specific evaluation items) and each of the specific evaluation items, as given in Table 6 in the report. (2,0,4)

S61 Maas and Owen (1973): over 12,000 student ratings of instructors and courses at Cornell University. The ranking in the present analysis is based on the size of the correlation between the overall rating of the instructor ("teaching skills of the teacher in comparison to other teachers") and each of 8 specific evaluation items concerning "Teaching and the Lecturer," as given in Table 1 in the report. (Note: 3 specific evaluation items were excluded because of differences in the nature of the rating scales in comparison to the 8 items that were used; see p. 4 of the report.) (0,0,1)

S62 Owen (1967): 798 students rating 16 instructors teaching 17 sections of introductory psychology at the University of Houston (Spring, 1966). The ranking in the present analysis is based on the size of the correlation between the overall rating of the instructor ("instructor's general teaching ability") and 24 specific evaluation items, as given in the Table on p. 186 in the report. Only the highest 24 of the 45 correlations have been ranked (3 items were tied for rank 25). (0,0,1)

S63 Plant and Sawrey (1970): 1,247 students (primarily undergraduates) rating 32 psychology instructors at a "large tax-supported state college in California." The ranking in the present analysis is based on the size of the correlation between the total rating score for the instructor (across 8 specific evaluation items) and each of the specific evaluation items, as given in Table 1 in the article. (0,1,0)

S64 Remmers (1929): student ratings of 115 instructors in as many different classes at Purdue University. The ranking in the present analysis is based on the size of the correlation between overall rating of the instructor (quintile placement of instructor when compared to other instructors) and each of 10 specific evaluation items, as given in the table on p. 20 in the article. (2,0,4)

S65 Remmers and Weisbrodt (1964): 1,908 students rating 59 instructors at Purdue University. The ranking in the present analysis is based on the size of the correlation of the overall rating of the instructor ("overall rating of the instructor") and each of 10 specific evaluation items, as given in Table 2 in the report. (2,0,4)

S66 Rosenshine, Cohen and Furst (1973): student ratings of instructors in 1,200 daytime, on-campus, undergraduate classes in all colleges and schools of Temple University (Spring, 1970). The ranking in the present analysis is based on the size of the correlation between the overall rating of the instructor ("instructor compared to other college instructors") and each of 19 specific evaluation items, as given in Table 3 in the article. (1,4,0)

S67 Sagen (1974): students rating 83 instructors at Morris Harvey College. The ranking in the present analysis is based on the size of the correlation between overall rating of the instructor ("overall instructional effectiveness") and each of 16 specific evaluation items, as given in Table 2 in the article. (0,1,1)

S68 Spencer (1967): 1,367 student ratings of instructors in the general engineering department of the University of Illinois. The ranking in the present analysis is based on the size of the correlation between the overall rating of the instructor ("overall effectiveness") and each of 11 specific evaluation items, as given in Table 1 in the report. (0,0,1)

S69 Van Horn (1968): 386 students rating instructors at Purdue University. The ranking in the present analysis is based on the size of the correlation between the overall rating of the instructor ("overall rating of instructor") and 12 specific evaluation items, as given in Appendix 2 of the report. (0,1,0)

S70 Walker (1968): 1,447 student ratings of 30 instructors at Lee Junior College (Spring, 1967). The ranking in the present analysis is based on the size of the correlation between the overall rating of the instructor ("overall rating of the instructor") and each of 10 specific evaluation items, as given in Table 9 in the report. (2,0,4)

S71 Widlak, McDaniel and Feldhusen (1973): student ratings of 208 instructors at Purdue University. The ranking in the present analysis is based on the size of the correlation between the overall rating of the instructor ("overall instructor") and each of 16 specific evaluation items, as given in Table 4 in the report. (0,1,0)

S72 Williams (1965): 777 juniors and seniors at three "private liberal arts colleges in the Twin Cities area" giving information about the best teacher and the poorest teacher whom they had had in college. The ranking in the present analysis is based on the degree to which each of 10 specific evaluation items ("instructional methods") discriminated the best instructors from the poorest ones, as given in Table 6:02 in the report. (0,0,0)

Each of the following studies contain information about the relationships between students' general evaluation of the teacher (or course) and their evaluation of specific attitudes or behaviors of the teacher, but none of these studies could be included as part of Table II for one or another reason (less than five specific evaluations are correlated with the indicator of overall evaluation; multi-item scales, rather than single items, are related to overall evaluation; the information given in the study is incomplete or the data are presented in such a way that the items of specific evaluation cannot be ranked): Aleamoni and Spencer (1973), Aleamoni and Yimmer (1973). Apt (196), Caffrey (1973), Cassel (1971), Corcoran (1961), Crannell

(1948), Deshpande, Webb, and Marks (1970), Dick (1967), Doyle (1972), Graham (1972), Granzin and Painter (1973), Hall (1970), Hildebrand, Wilson, and Dienst (1971), Kohlan (1973), Lahat-Mandelbaum and Kipnis (1973), Lathrop (1968), Levinthal (1974), Meinkoth (1971), Perlman (1973), Singhal and Stallings (n.d.), Solomon (1966), Spencer (1969), Sullivan and Skanes (1974), Voeks (1954-55), and Whitely, Doyle, and Hopkinson (1973). Studies by Bausell and Magoon (1972), Carter (1968), Hanke (1970), Pohlmann (1975), and Tobias and Hanlon (1974; cf. Tobias and Hanlon, 1975) were located too late to be included as part of Table II.

FOOTNOTES

¹Of the full set of 21 categories used throughout the present analysis, the two residual categories are not given in Table 1: Other professional characteristics (e.g., the instructor was punctual for class and for appointments; the instructor is highly accomplished in research; students willingly worked for this teacher; the instructor is active in campus activities) and personal characteristics (e.g., the instructor has a good sense of humor; the teacher was sincere and honest; the instructor is highly personable at all times in dress, voice, social grace, and manners; the instructor was free of personal peculiarities).

²A discussion of some of the advantages as well as the problems in this procedure can be found in Feldman (1971). Also in the interest of increased comparability, only studies presenting data for at least five characteristics were included in the analysis; moreover, only the 25 highest ranked characteristics were included in the calculation of the standardized ranks in those cases in which the list of characteristics in a study exceeded 25 (see Appendix A).

³Two kinds of studies are considered together here: (1) those showing preferences of students classified by their major fields and (2) those giving student preferences for teachers in different academic fields (regardless of the academic major of the student).

⁴If a study contained more than one item measuring the student's overall evaluation of the instructor, the most general or most global item was selected for analysis, and the others were dropped. Any item or items measuring the overall evaluation of the course was also excluded, except in the one case in which the overall evaluation of the course was the only available general measure (Jiobu and Pollis, 1971). The general measure used for two studies (Leftwich and Remmers, 1962; Plant and Sawry, 1970) was the total score across all specific items of the evaluation questionnaire, since neither study contained an item measuring overall evaluation of either the teacher or the course. Appendix B notes the indicator of general evaluation in each study that was used in the present analysis.

⁵The lists of instructor characteristics in studies of student preferences (analyzed earlier), as generated either by students themselves or by the researchers, focus almost exclusively on the qualities of the instructor and the instructional process rather than on the extent and nature of the benefit to students. That is, the focus is on the student's view of the process of instruction to the virtual exclusion of the student's view of the outcome of instruction (see Brandenburg, 1975; Hartley and Hogan, 1972; and Jaeger and Freijo, 1974). By contrast, some of the items in the teacher evaluation questionnaires in certain of the studies under analysis in this section, ask for students' judgments about the outcomes of the course for them in addition to their judgments about the instructional process itself. The following are

examples of such items: gaining of new knowledge was facilitated by the instructor; I developed significant skills in the field; I developed increased sensitivity and evaluative judgment; the instructor and the course have given me tools for attacking problems; the course has increased my general knowledge. Of the 23 studies on which Table II is based, five of them contain specific evaluation items codable into an outcome of instruction category (S50, S53, S54, S61, S66, in Appendix B). These items are generally highly correlated with overall rating of the instructor in these five studies: The overall average standardized ranking is .28. Had this outcome of instruction category been included in the presentation of data in Table III (col. 7), it would rank in third place, just after the categories of stimulation of interest in clarity and understandableness.

⁶Overall standardized ranks were also calculated for the categories of other professional characteristics and personal characteristics (see footnote 1), although they are not given in Tables I and II. None of these fractions were small enough to place in the top ten ranks in any of the seven columns of Table III.

⁷The rating procedure represented by the first two cases is not without its own problems and ambiguities, for example, those created by the incomparability of ratings made by raters with different value systems (Levinthal et al. 1971; Frey, 1974) or by raters who are using different frames of reference or reference points in evaluating an instructor (Grasha, 1975).

⁸These studies are Albino and Liberty (n.d.) Aleamoni and Spencer (1973), Apt (1966), Arreola (1973), Bejar and Doyle (1974), Bendig (1954), Brooks et al. (1971), Caffrey (1969, 1973), Carver and Liberty (1973), Cassell (1971), Centra (1973), Centra and Linn (1973), Coffman (1954), Cosgrove (1959), Costin (1971, 1974), Crannell (1953), Creager (1950), Davis (1969), Deshpande et al. (1970), Doyle (1972), Doyle and Liu (1972), Echandia (1963), Finkbeiner et al. (1973), French-Lazovik (1974; cf. French 1957), Frey (1973; cf. *Endeavor Information systems*, 1973), Fulcher and Anderson (1974), Gibb (1955), Hall (1970), Harley and Hogan (1972), Hildebrand et al. (1971), Hodgson (1958), Holmes (1971), Issacson et al. (1964), Jaeger and Freijo (1974), Kennedy (1975), Kohlan (1973), Leftwich and Remmers (1962), McKeachie (1972), Mann (1968), Meredith (1969), Owen (1967), Perry and Baumann (1973), Price and Magoon (1971). Purohit and Magoon (1971). Quereshi and Widlak (1973; cf. Widlak and Quereshi, 1972), Sherman and Blackburn (1975), Sockloff and Deaby (1971). Solomon (1966). Tobias and Hanlon (1974; cf. Tobias and Hanlon, 1975). Van Horn (1968). Veldman (1968). Villano (1975). Villano and Rosenstock (1973), Whitely et al. (1973), Widlak et al. (1973), and Yonge and Sassenrath (1968).

⁹Some examples of items coded as overall evaluation of the teacher or the course follow: compared to other instructors I have had, this instructor was excellent; the overall effectiveness of this instructor was high; I would definitely recommend this course to other students; this was one of the best courses that I have ever taken.

¹⁰The relatively weak association between the category of the nature and value of the course material and that of the overall evaluation of teacher that was found in the prior analysis of correlational studies, in contrast with the current connection between these two dimensions, may be partially explainable by the fact that the earlier analysis (unlike the one in this section) excluded items dealing with overall evaluation of the course (see footnote 4).