DOCUMENT RESUME

ED 140 877 JC 770 321

AUTHOR Bers, Trudy
TITLE An Analysis of the Student Evaluation Form at Oakton Community College.
PUB DATE Jun 77
NOTE 28p.

EDRS PRICE MF-$0.83 HC-$2.06 Plus Postage.
DESCRIPTORS Community Colleges; Correlation; *Evaluation Criteria; Factor Analysis; *Junior Colleges; *Measurement Instruments; Questionnaires; *Rating Scales; Reliability; *Statistical Analysis; Statistical Bias; Statistical Studies; *Student Evaluation of Teacher Performance; Validity
IDENTIFIERS Oakton Community College

ABSTRACT

This paper reports the results of an evaluative statistical analysis of the instrument employed for student evaluation of faculty at Oakton Community College (Illinois). The analysis was performed because the locally devised instrument had never been subjected to systematic study of reliability or validity, and because a review of pertinent literature indicated a lack of consensus on what types of variables affect students' evaluations of faculty. Results suggested that the instrument failed to discriminate clearly between positive and negative aspects of course organization, faculty performance, classroom ambience, and other evaluation variables; that the positive wording used in constructing the instrument was such that faulty data might result from inculcation of a response set; and that equal weighting was given to each item on the evaluation form, even though each item was not necessarily applicable to each instructor or class. It was recommended that Oakton define and clarify the purposes of faculty evaluation, review existing instruments in use at other institutions, select or design an instrument meeting the purposes, and pretest the chosen instrument for reliability and validity before implementation. A review of the literature, tabular data from the instrument analysis, a bibliography, and the evaluation instrument are included. (JDS)

AN ANALYSIS OF THE STUDENT EVALUATION FORM AT OAKTON COMMUNITY COLLEGE

Trudy Bers

June, 1977

The purpose of this report is to evaluate and analyze the student evaluation form currently in use at Oakton Community College. This form, which has been used since the spring semester of 1976, has never been subjected to systematic study. We have made many assumptions about it, and have included on it for the first time data about students as well as about student evaluations of faculty; however, these student data have never been used to crosscheck the overall means and standard deviations which have been the basis of interpretation.

The actual purpose of student evaluation of faculty has, I believe, been subject to question. Essentially evaluation can be for two distinct purposes: first, to discriminate among faculty on some predetermined criteria, so that faculty can be ranked, classified, or somehow rated against each other; or two, to provide student feedback for faculty in order that student-teacher relationships, classroom performance, and course organization by faculty members can be improved. The second purpose does not imply <u>rating</u> faculty or establishing standardized criteria. Because these purposes are so distinct, the evaluation form used in each system can be drawn very differently. Unfortunately, I think at Oakton we have neither clarified our purpose--we have talked of both, as though they were one--nor constructed an instrument to suit either objective. I do not believe this was intentional, and we have suffered frustration, anger, and hurt from our own best intentions. Nevertheless, since the college does seem committed to evaluation, and since the merit system will no longer raise red herrings in the way of evaluation,

my own view is that we can start afresh to first, clarify the purpose of our evaluation, and second, construct an evaluation process and instrument to achieve our purpose. This paper seeks to provide necessary background for these tasks.

The State of the Study of Student Evaluation

Hundreds of articles have been published about student evaluations of faculty. Among the more complete reviews of literature are those presented by Costin, et. al. (1971); Pasen (1977); and Shavelson and Dempsey-Atwood (1976). Generally they found that student evaluation forms can have a high degree of validity and reliability, and that mixed evidence exists about relationships among a variety of variables such as expected grade and teacher ratings (see below). The literature about student evaluations relies primarily on work done at four-year colleges and universities, although some work using community college students has been done (Ostrowski, 1975). This is disturbing, because the community colleges by design tend to put more emphasis on excellence in teaching than do senior institutions and because community college student bodies are very heterogeneous and may have differing standards from populations at four-year schools.

A major finding which consistently emerges from research on student evaluations is that evaluations are multidimensional. Two or more continua underlay student perceptions of teaching, and faculty who rank highly along one continuum may or may not rank highly along another one. While the number and definitions of continua implicit in any single evaluation instrument differ from one study to the next, certain similarities occur. Bolton, et. al. (1976) analyzed eleven separate

studies of student evaluation which together had isolated 75 factors. They concluded that these 75 factors could be reduced to six major dimensions; effort demanded by the instructor, instructor's preparation and organization, instructor's evaluation of student performance, instructor's knowledge of the subject, value of the course to the student, and instructor's friendliness and regard for students. They concluded that "the evaluation instrument which students complete should be scored within the major categories, generating a profile of subscale scores on the dimensions of teaching performance" (p. 119).

A second major concern evidenced by studies of student evaluations is that of the methodology employed. Essentially three variants of determining mean scores on ratings instruments are possible: within-class ratings, across-class ratings, and between-class ratings. Within-class ratings use each student evaluation as the unit of measurement; because all students are in the same class variables such as instructor-differences, subject matter-differences, and ambience within the classroom are held constant. However, because these are held constant differences between disciplines, teachers, times-of-day, etc., cannot be measured. A more serious problem for measuring within class means is that a large class is required. At Oakton the average number of students rating an instructor in a class is 10-15; this is too small a number to control for differences among students within the particular class. Pasen (1977) used nearly 500 students in a basic literature class for his within-class study.

Across-class ratings use as the unit of analysis each student response without regard for the class in which the evaluation is occuring. This is the easiest method to employ, but it results in the loss of

experience unique to a particular class. Whitely and Doyle (1976) recommend that across-class analysis not be used because it confounds within and between class correlations among variables. However, they themselves have used this approach (Doyle and Whitely,, 1974). Across-class ratings are used in the analysis of Oakton data presented below.

Between-class ratings used as the unit for analysis the mean scores of each class, rather than the scores of each evaluator. This approach has the advantage of reducing the impact of extreme evaluations, assumimg that they balance out within the class. It also accounts for unique class situations, since the class as a whole rather than the students within the class is the evaluator. Between-class ratings are also utilized in the data analysis below.

The abundance of studies about student evaluations noted above make a succinct summary of findings impossible. The contradictions within the literature make even an attempt at this frustrating. However, I have chosen to select out some findings which seem to me most germaine to the concerns enunciated by Oakton faculty and administration during our own informal analyses of the student evaluation instrument. The reader should remember that these are suggestive findings, not definitive ones, and the instrument, student sample, type of institution,, and timing of evaluation are among the the variables which may have affected these results.

Variables Affecting Evaluation

One of the questions most frequently posed about student evaluations is the affect of grades on ratings of instructors. Eagle (1977) found that student's expected grades were not correlated with their overall impression

of their courses, and Costin et. al. (1971), in an extensive literature review, suggest that there is mixed evidence about the correlation between actual grades received and evaluations. Pasen (1977) did find that within a single class there was an affect between expected grade and course evaluations. Centra (1977) and Doyle and Whitely (1974) used mean standardized final examination scores as a measure of course grades and found there was some relationship to various measures of instructor and course effectiveness. On the whole, then, it appears that evidence on all sides of this question can be found.

A second question often raised at Oakton has been whether the "entertainer" teacher is rewarded with good evaluations. Costin et. al. (1971) found no evidence of this, although Battle and Fabick (1975) did. However, studies consistently find that faculty enthusiasm is related to positive evaluations (Costin, et. al., 1971), and it may be that enthusiasm and the value of "entertaining" may spill over each other.

Students liking for the subject and interest in the course are related to positive evaluations (Doyle and Whitely, 1974), as is the student's major (Pasen, 1977). However, evidence about the affect of required and elective courses on evaluations is mixed (Costin, et. al., 1971).

In sum, then, it appears that whatever ideas one has about the variables affecting evaluations, one can find evidence to support those ideas. Thus it seems to me imperative that a strong effort be made to understand what variables affect evaluations at Oakton.

There is another set of concerns about what variables affect evaluations. These concerns operated in the realm of psychology and philosophy more than specific student characteristics, but they are, I think, important to note.

One of the unknowns with which we are operating at Oakton is what frames of reference or "anchors" (Pasen, 1977) students are using as they evaluate instructors. Students may well have internal criteria against which they measure their instructors. These internal criteria may be, for example, a student's best, worst, or average teacher (Gresha, 1975). That is, students may be implicitely measuring the instructor being evaluated against some other real or ideal teacher, and the choice of that referent may or may not affect evaluation outcomes. Follman, et. al. (1974) performed an experiment to test this and found the referent made little difference; however, they assigned referents to students. What we at Oakton do not know and have not asked is which frame of reference students are opting to use.

Another internal constraint which may be affecting evaluations is students' implicit theories about teaching characteristics which occur together (Whitely and Doyle, 1976). Students may have identified clusters of teacher behavior through past experiences, and in evaluating teachers assume that the occurrence of one such behavior is accompanied by the occurence of related behaviors, even if this does not occur in fact. For example, students may have implicit notions that a teacher who is enthusiastic about her subject is also responsive to students; thus, the enthusiastic teacher will be rated positively on responsiveness whether or not she is in fact responsive.

Finally, Morey, et. al. (1977) found that faculty have differing conceptions about what constitutes good teaching, as do students. Neither group, especially faculty, have reached consensus about this. Furthermore, they do not give the same importance, or weight, to the qualities which

are rated positively. Morey and his colleagues recommend that weighing items for importance before determining mean scores over items will provide a more accurate picture of evaluations.

As this cursory research note indicates, a variety of studies have been conducted in the general area of student evaluations of faculty, and a variety of conclusions have been reached. Probably the most pervasive conclusion of all is that there is a great deal about student evaluations that we do not know. This paucity of knowledge is important for several reasons. For those concerned about processes of learning and teaching, data about factors contributing to achievement and positive feelings toward learning (not necessarily the same thing) are crucial to the development of systematic learning theories and successful teaching techniques. For those concerned about accountability of educators--and I choose to view this term as a positive one--connections between the inputs to the educational system and the outputs of that system must be made. Meredith (1975) suggests three outcomes of instruction in higher education are identifiable: production, satisfaction, and growth. I think that viewing student evaluations as one tool for measuring outcomes of the educational system underscores the institutional need for such evaluations. Evaluations, as I view them, are a method for providing feedback to faculty; in turn faculty can use this information to upgrade their own teaching.

I have spent time reviewing methodological and philosophical issues of student evaluation and current findings for a specific reason: to illustrate the complexity of the subject, and the fact that while many researchers have carefully studied student evaluations, consensus about variables affecting ratings and even the dimensions along which ratings are assigned is not present. In one sense, then, the analysis below contributes to this confusion. In another, I hope it reduces confusion

for us at Oakton.

The Oakton Student Evaluation Form: An Analysis

The first step in evaluating the instrument was to determine correlations between each question (variable) on the instrument. (For exact wording of items and short titles used, see appendix I.) Pearson's correlation (referred to as Pearson's r, or simply r) was used as the statistic to measure congruence of responses for pairs of variables. The correlation varies from 1 to -1, with a positive score indicating that a high value on one variable is related to a high value on a second variable. A negative score indicates that a high value on one variable is related to a low value on the second variable. Pearson's r assumes that the realtionship between variables is linear (Garson, 1971).

A summary of means and standard deviations of all evaluative items is presented in Table 1. As both the low means and moderate standard deviations suggest, little discrimination among responses was obtained by the survey instrument. All means are strongly positive, and differences tend to be small. This should be borne in mind as the analysis continues: we are working essentially with minimal differences.

Three items on the instrument measured student inputs and expectations from the course. Nearly half the students indicate they prepare for class always or almost always, and another 86% say they do so often. Over 80% claim to attend always or almost always. Given the constant concern about low preparation and attendance evidenced by faculty, how can these conscientious respondents be explained? While the data do not lend themselves to interpretation, several speculations are plausible. One, the fall 1976 evaluations were given early in the term, before the noticeable drop the

last third of the term. Two, students answering the questionnaire are conscientious; but they are not representative of the entire population. Three, respondents are giving themselves more credit than is warranted. And four, a combination of these. Grade expectations also place students in a favorable light and illustrate grade inflation: nearly 35% of students expect an A, and over 80% expect an A or a B. Again, the early administration may have affected these expectations, but clearly students anticipate high grades.

Table 2 provides Pearson's correlations for all variables measured across-classes. Correlations among the items measuring faculty performance and course organization (questions 1 through 14) range from a moderate .273 between items 4 and 13, to a substantial .678 between items 8 and 9. Of more interest that these, however, is the correlation between variables measuring student input to a course and the evaluation of that course. Items 15 and 16 measure how often, by their own admission, students prepare for and attend class. The low correlations suggest that frequency of attendance and/or preparation for class are not related strongly to positive evaluations of the class. Item 17 measures the relationship between expected grade and class evaluation; once again, correlations are low. Expectations of good grades are not related to positive evaluations, although there is a relationship between frequency of preparation and expectations of high grades, not a surprising finding. Item 18 asks whether the course is elective or required for the student. Required courses were coded 1; elective courses were coded 2; and if the student did not know if the course was required, a value of 3 was assigned. Nearly 64% of the responses indicated the course was required, and one-quarter said it was elective. The remaining 14% either didn't know or had an invalid response. Given the

wide latitude students have in selecting courses beyond core requirements, it is likely that many students taking courses fulfilling general requirements (e.g., three credits of any social science) interpret this as absolute requirements. On the whole, all correlations are exceedingly small. There is little support for the assumption that students for whom a course is required evaluate the course differently than those for whom it is an elective.

Table 3 presents Pearson's correlations for between-class analysis. The same general pattern evidenced across-classes is present between-classes as well, although correlations are on the whole higher. There is a moderate relationship between how often the class as a whole prepares and evaluations for the course. Little evidence is available that courses which have a large number of students for whom the course is required are evaluated differently from courses which are for the most part elective.

Both Tables 2 and 3 provide support for the notion that positive evaluations of the course and instructor (items 1 through 14) are correlated with the student's recommending the course to others. (see coefficients in last row of each table). Course recommendations are not, however, strongly related to student's attendance, preparation, expected grade, or whether or not the course was required.

One of the frustrations in dealing with large numbers of Pearson's correlations as presented in the above two tables is the difficulty of interpreting patterns. As noted in the discussion of literature above, virtually all studies of student evaluations have determined that evaluation occurs along several continua which are both conceptually and empirically distinct from each other. In order to determine whether the Oakton evaluation distinguishes among two or more continua, a factor analysis on

across-class and on between-class evaluations was performed. This analysis determined whether two or more dimensions underlay the evaluation items 1 through 14 on the survey instrument. In other words, what items share a common underlying factor with what other items? (Garson, 1971).

For those knowledgable and interested in the factor analysis statistics, Tables 4 and 5 present factors extracted before rotation and factor loadings using both orthogonal and oblique rotations for across-class and between-class analysis, respectively. The most important finding, I believe, is that several distinct continuua do not underlay the Oakton evaluation items. Only two factors were extracted from the correlation matrix, and the first factor explains 89.7% of the explained variance in items in the across-class analysis and 89.2% of the explained variance in items in the between-class analysis. The second factor explains 10.3% and 10.8% of the explained variance, respectively.

The oblique rotation factor pattern, which presents the unique loadings of each item on the two variables, provides a clearer picture of items' relationship to each other than does the orthogonal factor matrix. It appears that items 1,2,10,11,12, and 13 load most highly on factor I. Items 5,6, and 7 load most highly on factor II. Items 8,9, and 14 load moderately on both factors. Items 3 and 4 load on factor II, but more moderately than do items 5,6, and 7. The higher the loading of a variable on a factor, the higher the proportion of variance in that variable explained by that factor.

Factor analysis was used as a method for teasing out of the Pearson's correlation matrix systematic patterns among evaluation items. In order to extend the analysis and at the same time to reduce the number of variables under consideration, I constructed two indices for use as dependent variables. The first index is the mean score on items 1,2, and 10 through 13. I have labeled this scale Course Cohesion, because it

draws from items relating to course organization, clarify, and the helpfulness of assignments in meeting course objectives. I have labeled the second scale Instructor Affect. It consists of the mean score on items 5,6, and 7, which measure instructor's responsiveness, sensitivity, and enthusiasm. Analysis below, then, uses these two scales as dependent variables.

Table 6 presents multiple correlations for across and between-class analysis, using Course Cohesion and Instructor Affect as dependent variables. Three independent variables were used in between-class analysis, and four independent variables were used across-class. All variables are measures of the amount of effort and/or concern students placed in the class. Whether a class is required was used only in the across-class analysis, because this is a variable relevant only to the individual student, not a measure of the class as a whole; a single course can be, for example, required for some students and an elective for others. The R^2 in each dependent variable section is the amount of variance in that dependent variable explained by the independent variables taken together. Thus course means for students' expected grades, preparation, attendance, and expected grade explain only 13% of the variance in across-class means on the index course cohesion. The moderate R^2s reaffirm that these student-related independent variables are not, on the whole, powerful explanations of varying evaluations (although they may be powerful for particular courses or faculty members).

Conclusions and Recommendations Regarding the Evaluation Instrument

I am concerned about the failure of the current student evaluation instrument to discriminate more clearly between positive and negative aspects of course organization, faculty performance, classroom ambience, etc. Further, I am concerned that despite our addition of student data to the instrument, we have failed to utilize these data. I suggest that these failings derive from several sources:

1. We have constructed the survey instrument so that all items are worded positively. Thus the response "all or almost all of the time" is always appropriate for positive evaluations, and never appropriate for critical ones. This can easily lead to a response set, a situation in which respondents are lulled into a pattern of response without their having to think very carefully about the particular item under consideration. This, in turn, provides faulty data.

2. We have weighted all items as equally important in calculating an overall course evaluation mean. Thus faculty who do not give tests, for example, but whom students persistently evaluate on item 13, are judged on this as well as more appropriate items.

3. We have not subjected our instrument to even rudimentary validity and reliability analysis. We have made assumptions about these qualities without testing for them. I believe that much of the skepticism about the evaluation form is rooted in distrust of it, and part of this distrust is legitimately rooted in the fact that we have not taken the time and effort to understand our instrument before using it.

4. We have not adequately used the data available on the form as it now stands. It would not be difficult to determine such simple facts as whether students taking a course because it is required evaluate their instructor differently from those who elected to take the course for each instructor interested in this. By failing to personalize data analysis we have prohibited instructors from taking maximum advantage of information available in raw form on the instrument.

5. We have consistenly constructed our own form without exploring the adoption or adaptation of existing forms. A number of student evaluation forms are available which have already been systematically tested for validity and reliability. We should, I believe, investigate these. If we

do wish to construct our own form again, I believe we should allow time and support for validating it.

Regarding the Evaluation Process

My comments in this section go somewhat beyond my analysis of the existing student evaluation form itself. They are founded in this analysis--and I think my above comments demonstrate I am critical of both the instrument and the use made of it--but draw as well upon my seven-year experience as an "evaluatee" at Oakton. Bluntly, I'm tired, and I think students who attend Oakton for any length of time are tired. I also think staff who are responsible for organizing evaluation are tired. We have been saturated with evaluation. Therefore, I suggest the following:

1. Declare a sabbatical on student evaluations. Except for new full and parttime faculty members (deans can work out criteria for defining "new"), faculty should not have courses evaluated until spring, 1978 at the earliest. This will provide a needed psychological and intellectual respite from evaluation saturation and time to accomplish my other suggestions.

2. Establish a special faculty or faculty/administration evaluation committee--ah yes, another committee. I envision the charge of this committee to be the following:

a. Collect and evaluate existing student evaluation forms in use at other institutions with a view of adopting one or more of these.

b. If none of these forms seems appropriate, adapt or construct a new one(s) for use at Oakton.

c. Plan and implement a program for pretesting the chosen instrument(s) for both validity and reliability.

d. Consider and make recommendations regarding

1. Form(s) to be used -- does every faculty member have to use the same form?
2. Frequency of evaluation -- does every faculty member have to be evaluated every term?
3. Timing of evaluation -- must every evaluation occur in the same week of the semester?

3. Once above recommendations and decisions have been made, appoint one appropriate individual to work out details for and implement the student evaluation process. This individual will probably need to work closely with the data processing staff to design and write programs which can analyze data in a form most useful to faculty members. (Incidentally, this is not really a difficult matter, although it will take time. The Statistical Package for the Social Sciences software package, which we have, can process a variety of data in an effective and flexible manner.)

Fundamental to all my recommendations are the following principles: we need to clarify the purposes of evaluation, be flexible in instrument and procedures, and be moderate in the number of evaluations we perform.

Table 1

Means and Standard Deviations
on Evaluative Questions

	Across-classes		Between-classes	
	$\bar{x}$	s.d.	$\bar{x}$	s.d.
Q1	1.533	.832	1.515	.410
Q2	1.441	.736	1.432	.335
Q3	1.601	.873	1.582	.416
Q4	1.815	.955	1.822	.476
Q5	1.388	.737	1.366	.333
Q6	1.472	.808	1.444	.379
Q7	1.420	.761	1.412	.386
Q8	1.754	.937	1.742	.506
Q9	1.654	.898	1.642	.502
Q10	1.606	.852	1.593	.399
Q11	1.767	.961	1.748	.430
Q12	1.686	.921	1.675	.397
Q13	1.965	1.122	2.061	.647
Q14	1.537	.836	1.530	.389

Table 2

Across-classes
Pearson's Correlations

Maximum N = 8034
actual N varies

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1																		
2	.6464																	
3	.4043	.3974																
4	.2957	.3079	.3606															
5	.4072	.3978	.4549	.4025														
6	.3866	.3875	.4338	.3943	.6592													
7	.4293	.3956	.5116	.3720	.5241	.5178												
8	.6022	.5511	.4915	.3611	.5192	.4803	.5335											
9	.5188	.5048	.4478	.3528	.4589	.4385	.5272	.6775										
10	.5138	.5841	.3759	.3168	.3693	.3743	.3684	.5459	.5244									
11	.4785	.4583	.3966	.3263	.3915	.3835	.3905	.5019	.4551	.5012								
12	.4827	.4582	.3731	.3069	.3581	.3582	.3671	.4629	.4190	.4919	.6303							
13	.4344	.3932	.3280	.2732	.3401	.3128	.3360	.4391	.3882	.4038	.4852	.5241						
14	.4495	.4467	.3505	.3211	.4397	.4618	.4013	.4889	.4234	.4437	.4221	.4292	.4794					
15	.2677	.2624	.2315	.1991	.1499	.1648	.1997	.2353	.1990	.2507	.2551	.2927	.2491	.2327				
16	.1779	.1737	.1634	.1209	.1174	.1213	.1469	.1422	.1350	.1547	.1384	.1795	.1336	.1536	.3690			
17	.2029	.1854	.1658	.1320	.1356	.1580	.1365	.1962	.0908	.1660	.1627	.1825	.1877	.2513	.4089	.2034		
18	.0260	.0195	-.0065	-.0011	-.0207	-.0175	-.0188	-.0058	.0005	.0263	.0258	.0402	.0418	.0047	.0843	.0793	.0436	
19	.4275	.3862	.3765	.2819	.4411	.4266	.4340	.5111	.4615	.3407	.3487	.3390	.3152	.3980	.1224	.1114	.1377	-.0018

N = 514

Table 3

Between-classes
Pearson's Correlations

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1																		
2	.8071																	
3	.5562	.5490																
4	.3296	.3487	.4464															
5	.5545	.5746	.6121	.4920														
6	.5091	.5582	.5432	.4338	.8221													
7	.5883	.5705	.7285	.4681	.6535	.6323												
8	.7761	.7424	.6752	.4343	.6941	.6095	.6744											
9	.7082	.7396	.6065	.3866	.5887	.5559	.6566	.8203										
10	.6987	.7826	.5179	.3430	.4966	.4817	.4954	.6775	.6908									
11	.6571	.6929	.5592	.3510	.5369	.5053	.5529	.6662	.6394	.6846								
12	.6763	.6701	.5082	.3013	.4513	.4174	.5121	.5956	.5870	.6438	.7184							
13	.4591	.4381	.2017	.1212	.2109	.1839	.2740	.3718	.3996	.4327	.4618	.4605						
14	.5645	.6112	.4535	.3266	.5532	.5757	.5119	.6001	.5408	.5825	.5714	.5341	.4277					
15	.3647	.3580	.4006	.2317	.1860	.2162	.2983	.3570	.2775	.3389	.3806	.4014	.2395	.3218				
16	.2491	.2622	.3255	.1509	.1136	.1451	.2290	.1906	.2182	.2150	.2069	.1844	.0863	.1787	.4095			
17	.2225	.2384	.3235	.2488	.2465	.3030	.2551	.3134	.1540	.1982	.2146	.1849	.0378	.4167	.4886	.2149		
18	-.0201	-.0497	-.1746	-.1327	-.2008	-.1762	-.1709	-.1513	-.1131	-.0604	-.0840	-.0232	-.0903	-.0716	.0180	.0506	-.1017	
19	.6700	.6546	.6486	.3790	.7235	.6549	.6606	.7790	.7204	.5433	.5902	.5184	.3245	.5710	.1938	.1631	.2452	-.1712

Table 4

Across-classes
Factor Analysis

FACTOR	EIGENVALUE	PCT OF VAR	CUM PCT
1	6.33304	89.7	89.7
2	0.72704	10.3	100.0

	Orthogonal Varimax Factor Matrix		Oblique Factor Pattern (unique loadings)	
	Factor 1	Factor 2	Factor 1	Factor 2
Q1	0.67374	0.33190	0.71858	0.04429
Q2	0.66958	0.30176	0.72935	0.00705
Q3	0.36809	0.52123	0.20860	0.47142
Q4	0.28231	0.46038	0.12716	0.44133
Q5	0.25075	0.75623	-0.07505	0.84897
Q6	0.24170	0.73389	-0.07502	0.82485
Q7	0.32254	0.64287	0.08202	0.65800
Q8	0.58607	0.52233	0.49862	0.34590
Q9	0.52898	0.48589	0.44222	0.33122
Q10	0.67458	0.27545	0.75025	-0.03047
Q11	0.65388	0.29704	0.71096	0.00999
Q12	0.67914	0.23611	0.76802	-0.08070
Q13	0.58020	0.26073	0.63239	0.00513
Q14	0.49417	0.41677	0.43324	0.26054

Factor correlations

	I	II
I	1.000	.71904
II	.71904	1.000

Table 5

Between-courses
Factor Analysis

FACTOR	EIGENVALUE	PCT OF VAR	CUM PCT
1	7.88374	89.2	89.2
2	0.95564	10.8	100.0

	Orthogonal Varimax Factor Matrix		Oblique Factor Pattern (unique loadings)	
	Factor 1	Factor 2	Factor 1	Factor 2
Q1	0.74985	0.42129	0.73800	0.17696
Q2	0.77086	0.43321	0.75863	0.18205
Q3	0.36958	0.66550	0.17499	0.69919
Q4	0.17467	0.52894	-0.00569	0.56058
Q5	0.25864	0.84842	-0.03531	0.90868
Q6	0.25001	0.78638	-0.02023	0.83770
Q7	0.37098	0.72055	0.15398	0.70495
Q8	0.61558	0.63146	0.48810	0.48960
Q9	0.63780	0.53698	0.55407	0.36589
Q10	0.75219	0.35804	0.76692	0.09968
Q11	0.70737	0.40527	0.69296	0.17639
Q12	0.72817	0.30756	0.75853	0.04942
Q13	0.59019	0.05329	0.69559	-0.19623
Q14	0.53752	0.25519	0.46587	0.31154

Factor correlations

	I	II
I	1.000	.62712
II	.62712	1.000

Table 6

Multiple Correlations: Course Cohesion and Instructor Affect explained by Student Input

	Across-classes[1]		Between-classes[2]	
	R	R^2	R	R^2
Course Cohesion	.36	.13	.42	.18
Instructor Affect	.23	.05	.34	.11

[1] Independent variables include required or elective course, expected grade, frequency of preparation, and frequency of attendance. Students who did not know whether a course was required were excluded from analysis.

[2] Independent variables include course averages on expected grade, frequency of preparation, and frequency of attendance.

REFERENCES

Battle, James, and Fabick, Stephen. 1975. "Validity of College Students' Evaluations of Instructors' Competence," Psychological Reports. 37:1112-1114.

Bolton, Brian, Bonge, Dennis, and Hinman, Suki. 1976. "Dimensions of Students' Evaluation of Instruction: An Empirical Synthesis," Psychological Reports. 38:119-123.
1977

Centra, John A. "Student Ratings of Instruction and Their Relationship to Student Learning," American Educational Research Journal, 14:17-24.

Costin, Frank, Greenough, William T., and Menges, Robert J. 1971. "Student Ratings of College Teaching: Reliability, Validity, and Usefulness," Review of Educational Research. 41:511-535.

Doyle, Kenneth O., and Whitely, Susan E. 1974. "Student Ratings as Criteria for Effective Teaching," American Educational Research Journal. 11:359-274.

Eagle, Norman. 1977. "Studies in the Reliability and Validity of a Community College Rating of Instruction Questionnaire," Community/Junior College Research Quarterly. 1:303-314.

Follman, John, Lavely, Carolyn, Silverman, Stuart, and Merica, John. 1974. "Student Raters' Referents in Rating College Teaching Effectiveness," The Journal of Psychology. 86:247-249.

Garson, G. David. 1971. Handbook of Political Science Methods. Boston: Holbrook Press, Inc.

Meredith, Gerald M. 1975. "Toward a Systems Approach to Student-Based Ratings of Instruction," The Journal of Psychology. 91:235-246.

Morey, T. Mark, Garthwaite, Roger A., and Zimowki, Michele F. 1977. "Teaching Effectiveness: Actual, Ideal, and Importance Ratings," Unpublished paper, State University of New York, College at Oswego.

Ostrowski, M.V. 1975. "A Comparison of Grades Students Achieve at William Rainey Harper College and How They Rate the Effectiveness of Their Instructor at Mid-Term During the Spring 1975 Semester," Unpublished paper.

Pasen, Robert M. 1977. "The Differential Effect of Grade, Sex, and Discipline on Two Global Factors: A Within-Class Analysis of Student Ratings of Instruction," Unpublished Ph.D. dissertation, Northwestern University.

Shavelson, Richard, and Dempsey-Atwood, Nancy. 1976. "Generalizability of Measures of Teaching Behavior," Review of Educational Research. 46: 553-611.

Whitely, Susan E., and Doyle, Kenneth O. 1976. "Implicit Theories in Student Ratings," American Educational Research Journal. 13:241-253.

STUDENT EVALUATION OF FACULTY

Student Signature (optional) PART I IDENTIFIER

This is an evaluation of your instructor and the course. Please give serious consideration to the survey.
The instructor will use the information to assess his/her effectiveness. The administration will use the information as part of the total evaluation of faculty, course, and programs.

Use no. 2 pencil.

OPEN END COMMENTS

Comment on how well the exams, quizzes, and/or projects contribute to your learning (e.g. difficulty, fairness, appropriateness, etc.).

Comment on the value of books, out-of-class assignments, papers, labs, or projects in this course.

What did you like most about this course/instructor?

From what classroom activities did you learn best in this course?

What did you not like about this course/instructor?

From what classroom activities did you learn least in this course?

List any specific recommendations you have for the instructor.

Turn over for Part II

STUDENT EVALUATION OF FACULTY

PART II

Please choose the words that fit for you. Mark IBM card.

IF THE STATEMENT DOES NOT APPLY TO THIS COURSE, LEAVE THE SPACE BLANK.

a. Always or almost always
b. Often
c. Some of the time
d. Seldom
e. Never or almost never

1. You know what you are supposed to be learning in this course.
2. You know what you are supposed to do for this course.
3. The instructor encourages attentiveness and/or participation.
4. The instructor is available outside of the class.
5. The instructor is responsive to student's comments and/or questions.
6. The instructor is sensitive to and respectful of his/her students.
7. The instructor is enthusiastic about the course.
8. The instructor presents subject matter clearly.
9. The instructor is well-organized.
10. Reading and writing assignments are clear.
11. Follow up to reading/writing assignments is helpful to your learning.
12. Assignments are helpful in your learning.
13. Examinations are helpful in your learning.
14. Grading is fair.
15. You consider yourself a person who prepares for this course (e.g. class participation, examinations, reading assignments, papers, etc.
16. You attend this class.
17. Mark on the IBM card the small letter that corresponds with the grade you expect in this course.

    a. A b. B c. C d. D e. R

18. This course is required in your chosen curriculum. a. Yes b. No c. Don't know
19. You would recommend this instructor to a friend. a. Yes b. No