

Winter 1995

Factors Underlying Effective College Teaching: What Students Tell Us

Carolyn Ridenour

University of Dayton, cridenour1@udayton.edu

Stephen J. Blatt

University of Dayton, blattste@udayton.edu

Follow this and additional works at: https://ecommons.udayton.edu/eda_fac_pub



Part of the [Educational Assessment, Evaluation, and Research Commons](#), [Educational Leadership Commons](#), and the [Scholarship of Teaching and Learning Commons](#)

eCommons Citation

Ridenour, Carolyn and Blatt, Stephen J., "Factors Underlying Effective College Teaching: What Students Tell Us" (1995). *Educational Leadership Faculty Publications*. 84.

https://ecommons.udayton.edu/eda_fac_pub/84

This Article is brought to you for free and open access by the Department of Educational Leadership at eCommons. It has been accepted for inclusion in Educational Leadership Faculty Publications by an authorized administrator of eCommons. For more information, please contact frice1@udayton.edu, mschlangen1@udayton.edu.

Factors Underlying Effective College Teaching: What Students Tell Us

Carolyn Benz and Stephen J. Blatt, University of Dayton

THESE MATERIALS PROVIDED BY
BALL STATE UNIVERSITY LIBRARIES
NOTICE: This material may be protected
by Copyright Law (Title 17, US Code)

Abstract

The researchers analyzed 28,000 student evaluations of faculty across 46 departments for one academic term. A 27-item instrument on which students rated faculty was used. One global item assessing overall instructor effectiveness was predicted most strongly by three items: namely, students' perception that the instructor was prepared, presented subject matter clearly, and was interesting. The predictors of students' perceiving that they "learned a lot" were the ratings on three items: the instructor was interesting, the course met the objectives, and the instructor was well-prepared. Being prepared and being interesting seem to be critical characteristics for university faculty in the classroom.

Other than the routine reports for individual professors there had never been a systematic study of the student evaluations of faculty at our institution. We decided to examine the aggregate of evaluations for one semester in ways that could reveal the underlying dimensions of student ratings. This became our first line of inquiry. Results would add to faculty's understanding of the validity of a measuring tool that has been used for many years, and one that impacts seriously on their pay and promotion. A common administrative use of only one item ("Overall rating of the instructor"), to the exclusion of 26 other items on which data were routinely collected, posed a second intriguing question for us; namely, Which of the other items weighted most heavily in predicting the rating of "Overall effectiveness" of the instructor? And, thirdly, Which of the items were the most important predictors of students' reporting that they "learned a lot?"

Research on student evaluation of faculty seems to adopt one of two perspectives: either teaching effectiveness can be assessed globally, using a single overall measure; or teaching is multidimensional and assessment must address many individual dimensions (Blatt & Benz, 1993; Ryan, Harrison, & Zia, 1993). Ryan and colleagues conducted an extensive review of published studies and found a lack of research that examined individual teaching behaviors that relate to a one-item global evaluation. While a void exists for that particular type of study, the field in general is well-researched. Marsh and Bailey (1993) report that literally thousands of studies have been conducted and they concluded that the process itself seems supportable in assessing teaching effectiveness.

Procedures

At the conclusion of each term at the University of Dayton undergraduate students are asked to fill out a formal evaluation form in each course. In addition to demographic items (gender, year in school, GPA, whether the course is required or not, etc.) there are items that relate to quality of the instruction and the course. The students respond to the items anonymously on bubble-scan sheets which

are collected by one member of the class and sent to the computer center. Results for each course are returned to each instructor from the computer center, via each department chair.

The data from student evaluations of the January-April 1992 term were aggregated and analyzed. Data from 46 departments, university-wide, were analyzed separately. All were then aggregated to form a data-set of approximately 28,800. There was one limitation in the design of the study: the analyses violated the assumption of independence of measures. There were not 28,800 separate student evaluations. A student would typically enroll in four or five courses, therefore, completing four or five evaluation forms at the end of the term. Because the students complete the evaluation forms anonymously there was no way to correct for non-independence of the data.

Results

Question #1: Underlying dimensions of the evaluation instrument

A factor analysis of the data was conducted to reveal the underlying structure of the evaluation instrument. A principal components solution with varimax rotation, with an eigenvalue cutoff of 1.00, was used. Selecting this type of factor analysis grew from the traditional notion of factor analysis as a way to map an unfamiliar terrain, as Rummel (1970) puts it, and our desire to reveal the clearest and simplest structure underlying faculty evaluation, i.e., uncorrelated factors.

A three-factor solution resulted and Table 1 reports the results. The first factor defines a dimension interpreted as "Instruction." Items originally designed to relate to both the instructor and the course loaded on this factor. This factor includes the global or "overall" items that call for assessing the course and assessing the instructor (items #7 and #8).

Factor Two was interpreted to be an "Affective" dimension. Four items addressing such issues as student's expressiveness, instructor's willingness to help, etc. loaded

highly on this factor. The third factor, named "Materials", was interpreted as a dimension clearly related to materials and scheduling.

In sum, these results would suggest that the items on the evaluation instrument group around three dimensions: "Instruction", "Affect", and "Materials."

Question #2: Predictors of overall instructor rating

Item #8 states: "Everything considered, how would you rate this instructor?" This global evaluation item is the most important item for faculty concern. In some departments this item is the exclusive means of evaluation. We were interested in which specific teaching behaviors and attitudes related most strongly to it. In order to answer this question a regression analysis was done. The items categorized by the university as "instructor items" (#10 through #17) were used as predictor variables against the criterion of item #8: Overall instructor rating. These items accounted for .65 of the variance of the overall instructor perception; the regression weights appear on Table 2. Tentatively, the three items contributing most to the perceptions of the instructor were item 10 "The instructor was prepared well", item #12, "The subject matter was clearly presented...", and item #13 "The instructor put material across in an interesting way." The other teaching behavior and attitude items contributed little to the overall ratings of instructor compared to these three items.

Because regression weights are unstable from sample to sample and one can conclude little from just one data-set, a cross-validation study was done. The results supported the pattern of weights. We have confidence the three items are, in fact, significant characteristics related to overall instructor ratings.

Question #3: Predictors of students' perception of "having learned a lot?"

Item #9: "I learned a great deal from this course" is all but forgotten in faculty evaluations at this institution. Responses to this item seemed to us to be the most relevant one of all as far as our goals as faculty are concerned. A legitimate case could be made for the fact that students reporting a sense of having learned a lot is even more powerful effectiveness indicator than a global assessment item like #7 or #8. That a student may rate an instructor in a less than positive way "overall" is relatively unimportant, one might assert, compared to whether or not the student reported having learned.

This item, item #9, is not part of the "overall" ratings of course and instructor. It is never used in the typical departmental review of faculty.

In order to determine which items on the instrument were most predictive of students' having reported a feeling of learning a lot, all the items (#10 through #25) were regressed on item #9 as a criterion variable. Table 3 presents the results.

The 16 items accounted for .53 of the variance in the students' reported sense of learning. The most contributory items were #13: "The instructor put material across in

an interesting way", item #19: "The course effectively met these objectives", and item #10: "The instructor was prepared well for classes."

Three items were negatively related to the criterion, #11: "The instructor spoke clearly and audibly", #16: "The instructor respected students as persons" #25: "Examinations and assignments were graded and returned within a reasonable time."

Cross-validation studies supported this pattern of importance among the variables.

Discussion

This study of student evaluations at our institution was born out of faculty self-interest; promotion and pay are strongly linked to the student evaluation system. We felt that to assist ourselves and our colleagues to become acquainted with the student evaluation process, we would examine the data generated from our own students for insights into the constructs being measured and how the global (or "overall") items relate to items on specific teacher behaviors and attitudes. In many departments, these "overall" items, #7 and #8, are used to the exclusion of all other items. Most frequently, as a matter of fact, only #8 is used. Additionally, we wanted to explore the relationships between the other items and item #9, the students' perception of having learned a lot.

We drew several conclusions from the data analysis and presented them to the faculty during an Inservice-Day shortly after the research was completed. Discussion begun during that meeting continued informally for several weeks after.

The conclusions are clear. First of all, students seem to be attending to behavioral factors rather than affective factors in their overall evaluation of faculty (item #8). This conclusion is warranted by our finding that the following items were predictive of overall instructor rating:

The instructor prepared well

The subject matter was clearly presented

The instructor put material across in an interesting way.

On the other hand, the items more reflective of "affect" were not strongly predictive of instructor ratings. Examples of these items are those that address "respect for students" and "fairness." This is particularly meaningful because departments often use only item #8 for personnel decisions. It may well be that being "nice" and supportive with students is not a sure pathway toward high teaching ratings. Students may be telling us "Be interesting and prepared; niceness won't cut it!"

Secondly, item #9: "I learned a lot" provided an interesting parallel to item #8. Our perception that this item is mostly ignored was confirmed when we discovered that a number of faculty had forgotten it was even on the form. The responses to this item may be more important than responses to item #8 where the instructor overall is rated. That students' perceptions about learning may be solid evidence of effective teaching. We found that students' feelings that

Table 1
Factor Analysis Results: Student Evaluation of Faculty Form
(Data from January 1992 term)

Item	Question	Instruction	Affective	Materials
7	Everything considered, how would you rate this course?	.74443		
8	Everything considered, how would you rate this instructor?	.76032		
9	I learned a great deal from this course.	.73990		
10	The instructor prepared well for classes.	.70493		
11	The instructor spoke clearly and audibly.	.56769		
12	The subject matter was clearly presented by the instructor.	.75327		
13	The instructor put material across in an interesting way.	.71228		
14	Students were able to express themselves freely as a result of the instructor's openness to their ideas.		.76411	
15	The instructor was willing to help students who experienced difficulty in the course.		.73469	
16	The instructor respected students as persons.		.79901	
17	The instructor was fair in grading examinations and assignments.		.56227	
18	The goals and objectives of this course well defined.	.66159		
19	This course effectively met these objectives.	.70905		
20	This course was well coordinated and well organized.	.74749		
21	Supplemental course material, such as handouts, visual aids, bibliographies, etc., enriched this course.	.47011		
22	The textbook was an asset to this course.			.79282
23	Assignments were relevant to course content.			.63402
24	Examinations related well to the material emphasized in the course.			.55747
25	Examinations and assignments were graded and returned within a reasonable time to students.			.44881
	Eigenvalue	5.970	3.328	2.674
	Trace Variance	.31	.17	.14

{62% trace variance}

Table 2
Items #10 through #17 as Predictor
Variables of Overall Rating of Instructor (#8)

Item	Question	partial regression weight	t	p>t
10	The instructor prepared well for classes.	.2110	36.35	.0001
		.2332 (1)	28.16 (1)	
		.1906 (2)	23.39 (2)	
11	The instructor spoke clearly and audibly.	.0256	4.51	.0001
		.0295 (1)	3.46 (1)	
		.0238 (2)	3.102 (2)	
12	The subject matter was clearly presented by the instructor.	.2267	41.01	.0001
		.2237 (1)	28.67 (1)	
		.2285 (2)	29.16 (2)	
13	The instructor put material across in an interesting way.	.2851	59.80	.0001
		.2781 (1)	41.85 (1)	
		.2913 (20)	42.52 (2)	
14	Students were able to express themselves freely as a result of the instructor's openness to their ideas.	.0389	7.218	.0001
		.0473 (1)	6.17 (1)	
		.0321 (2)	4.236 (2)	
15	The instructor was willing to help students who experienced difficulty in the course.	.0619	11.481	.0001
		.0612 (1)	8.04 (1)	
		.0643 (2)	8.387 (2)	
16	The instructor respected students as persons.	.0977	14.95	.0001
		.0859 (1)	9.301 (1)	
		.1080 (2)	11.66 (2)	
17	The instructor was fair in grading examinations and assignments.	.0982	20.743	.0001
		.1056 (1)	16.31 (1)	
		.0885 (2)	12.71 (2)	

Note: $R^2 = .6546$ $df = 8/27734$ R^2 cross validation = .6392

Table 3

All items (#10-#25) as predictor variables of students' perceptions of "learning a lot" (#9)

Item	Question	partial regression weight	t	t>p
10	The instructor prepared well for classes.	.1598 .1721 (1) .1489 (2)	21.53 15.87 (1) 14.63 (2)	.0001
11	The instructor spoke clearly and audibly.	-.0347 -.0332 (1) -.0366 (2)	-5.20 -3.27 (1) -4.112 (2)	.0001
12	The subject matter was clearly presented by the instructor.	.0910 .0712 (1) .1115 (2)	13.32 7.31 (1) 11.62 (2)	.0001
13	The instructor put material across in an interesting way.	.2150 .2400 (1) .1898 (2)	37.62 29.67 (1) 23.47 (2)	
14	Students were able to express themselves freely as a result of the instructor's openness to their ideas.	.0058 -.0121 (1) .0225 (2)	0.92 -1.33 (1) 2.52 (2)	.3575
15	The instructor was willing to help students who experienced difficulty in the course.	.0109 -.0007 (1) .0234 (2)	1.72 -.077 (1) 2.63 (2)	.0848
16	The instructor respected students as persons.	-.0151 -.0169 (1) -.0117 (2)	-1.98 -1.54 (1) -1.10 (2)	.0472
17	The instructor was fair in grading examinations and assignments.	.0223 .0281 (1) .0145	3.82 3.49 (1) 1.71 (2)	.0001
18	The goals and objectives of this course were well defined.	.0079 -.0067 (1) .0213 (2)	0.97 -0.56 (1) 1.88 (2)	.3297
19	This course effectively met its objectives.	.2209 .2218 (1) .2184 (2)	24.89 16.84 (1) 18.17 (2)	.0001
20	This course was well coordinated and well organized.	.0736 .0850 (1) .0612 (2)	9.40 7.53 (1) 5.63 (2)	.0001
21	Supplemental course material, such as handouts, visual aids, bibliographies, etc., enriched this course.	.0574 .0424 (1) .0736 (2)	11.05 5.84 (1) 9.86 (2)	.0001
22	The textbook was an asset to this course.	.0460 .0412 (1) .0499 (2)	11.68 7.52 (1) 8.76 (2)	.0001
23	Assignments were relevant to course content.	.0517 .0748 (1) .0333 (2)	7.98 7.62 (1) 3.85 (2)	.0001
24	Examinations related well to the material emphasized in the course.	.0488 .0485 (1) .0479 (2)	8.35 5.81 (1) 5.84 (2)	.0001
25	Examinations and assignments were graded and returned within a reasonable time to students.	-.0328 -.0288 (1) -.0392 (2)	-5.99 -3.73 (1) -5.00 (2)	.0001

Note: $R^2 = .5380$ $df = 16/25660$ R^2 cross validation = .5385

*the instructor prepared well
the instructor put material across in an
interesting way*

the course effectively met these objectives

were predictive of student's perception of "having learned." The first two of these behaviors are the same behaviors that predicted responses to item #8. Clearly, being interesting and having prepared are crucial for both being perceived as a good teacher and for students' feelings of learning.

Beyond the results reported here, it is interesting that we found no gender difference in the evaluation of faculty; however, female students rated faculty characteristics higher on all items than did male students. Whether the student viewed the course as required or not required made no significant difference in their responses to course evaluation. Surprisingly, the instructor rating/course rating, items #7 and #8, and the "amount student learned" item, #9, had higher means when the course was not required.

Finally, we return to our original interest - the underlying dimensions of what is being measured. The student evaluation form appears to be measuring "instruction", "affect", and "materials." The two "overall" ratings (#7 and #8) loaded heavily on the "instruction" factor. This suggests that administrators can use these items with confidence in assessing overall faculty proficiency in instruction. The underlying dimensions of "instruction" and "course" overlap, according to our factor analysis, evidence that in students' minds the two are inextricably linked. While this raises other questions, (Does subject matter preference bias student ratings?), it seems to confirm that what one perceives about the course is also likely to be what one perceives about the instructor.

The need for systematic institutional studies

Following discussion of these findings with faculty during In-Service day, findings were also printed in the campus newspaper. Feedback from faculty overwhelmingly supported this line of inquiry. They have asked for more study of the instrument and student responses, and also of the ways in which the results are used from department to department. Many professors raised issues of specific relevance; i.e., Is a universal form the best tool? For example, do laboratory courses and performance courses in the fine arts present a different set of dynamics for students to assess than classroom lecture courses? Some of these same faculty suggested an additional response option of "Does not apply." Some qualities queried on the form were considered irrelevant to some courses.

A few faculty spoke to a need for opportunity for faculty feedback to student evaluations. When or how do faculty have a voice in their use? Others were interested in whether or not students' evaluations were related to course grades they received. Some recommended open-ended response options for students. While some departments have added room for comments under each item, many faculty pressed for requiring students to write reasons why they gave the numerical rating that they did. The possible relationship between students' personal investment in the course to

how they evaluate it was discussed. In other words, might there be value in asking students how many classes they missed and how much study-time they put in per week on the course material? Also, some asked for inclusion of items on gender-sensitive and ethnicity-sensitive language and attitudes on the part of faculty.

Finally, the order of items on the Student Evaluation Form has been questioned. The overall, global, assessment items are currently #7 and #8 and, as such, precede the items on specific teaching behaviors and attitudes. Whether or not this order encourages the appropriate response-set among students is a concern. In the past, these two global items were at the end of the list of all other items. The resulting dynamic of both strategies needs to be assessed, according to some faculty.

Our intent is to continue this research in a variety of ways. The immediate plan is to ask students to record the "meaning" of their numerical responses in a randomly selected set of classrooms. In other words, what does it mean to students to: "put material across in an interesting way?" Further policy studies of administrative use of the process are also planned. A more systematic study of faculty views is required. We plan to interview faculty, as well as students, to get more in-depth interpretations of this process. How do faculty relate to the process personally and professionally? What impact does it have or not have on their teaching?

At the present time a replication of this study as well as an examination of the communication competencies that correlate with these items currently underway. We strongly believe that continuing broad-based institutional examination of the process of student evaluation is absolutely essential for a positive climate of optimal teaching and learning at our university to flourish.

References

- Blatt, S. J. & Benz, C. R. (1993) The relationship between communication competence and perceived faculty effectiveness. Paper presented at the annual meeting of the Southern States/Central States Communication Association, Lexington, KY.
- Marsh, H. W. & Bailey, M. (1993) Multidimensional students' evaluations of teaching effectiveness: A profile analysis. *Journal of Higher Education*, 64(1), 1-18.
- Rummel, R. J. (1970) *Applied Factor Analysis*. Northwestern University Press, Evanston, Illinois: North Western University.
- Ryan, J.M., Harrison, P.D. & Zia, Yi-Mei. (1993) The relationship between individual instructional characteristics and the global assessment of teaching effectiveness across different instructional contexts. Paper presented at the annual meeting of the American Educational Research Association, Atlanta, GA.