DOCUMENT RESUME

ED 078 748 HE 004 296

AUTHOR Vogt, Karl E.; Lasher, Harry
TITLE Does Student Evaluation Stimulate Improved Teaching?
INSTITUTION Bowling Green State Univ., Ohio. Coll. of Business Administration.
PUB DATE [73]
NOTE 14p.

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS College Faculty; College Students; *Effective Teaching; *Higher Education; Student Attitudes; *Student Opinion; Student Role; *Teacher Evaluation; *Teacher Improvement

ABSTRACT

Student appraisal of faculty instructional competencies is commonplace. Although several logics account for sponsorship of student evaluation schemes, the ultimate product of student evaluation ought to be improved instruction. The purpose of this paper is to investigate relationships between student evaluating and better teaching. A mandatory system of student assessment of teaching skills employed at Bowling Green University is the frame of reference. If student ratings contribute to better teaching, ratings should improve over time. Regression equations and standards tests were employed to determine the existence of trend increments. Findings reveal that regression coefficients of regression equations were as low as 0. By inference, student evaluation had not contributed to better teaching. Shortcomings in the administration of the evaluation scheme and faculty attitudes and capabilities account for apparent failures of the scheme to result in improved teaching. Appendices include related research material. (Author)

DOES STUDENT EVALUATION STIMULATE IMPROVED TEACHING?

Karl E. Vogt and Larry Lasher

ABSTRACT

Student appraisal of faculty instructional competencies is commonplace. Although several logics account for sponsorship of student evaluation schemes, the ultimate product of student evaluation ought to be improved instruction. The purpose of this paper is to investigate relationships between student evaluating and better teaching. A mandatory system of student assessment of teaching skills employed at Bowling Green University is the frame of reference.

If student ratings contribute to better teaching, ratings should improve over time. Regression equations and standard t tests were employed to determine the existence of trend increments. Findings reveal that regression coefficients of regression equations were as low as 0. By inference, student evaluation had not contributed to better teaching.

Shortcomings in the administration of the evaluation scheme and faculty attitudes and capabilities account for apparent failures of the scheme to result in improved teaching.

DOES STUDENT EVALUATION STIMULATE IMPROVED TEACHING?

Over the past several years, student evaluation of faculty instructional competencies has become fashinable in academe. A variety of factors account for institutional acceptance of the student evaluation concept. Evaluation schemes may be viewed as a device to provide faculty insights about their abilities to help students learn or the purpose of student evaluation may be tied to the reward system--that is, evaluation results are used to document claims for additional benefits. Regardless of the motives for acceptance, we would, nevertheless, expect the ultimate product of student evaluation to be an improvement in teaching effectiveness. This is to say that the real test of student evaluation has to do with an improvement in teaching effectiveness over time.

The purpose of this paper is to gauge the impact of student assessments of faculty instructional skills on the improvement of teaching effectiveness (as inferred from student evaluation results over time). Research draws upon experiences and empirical data associated with student evaluation at the College of Business Administration, Bowling Green State University.

I

THE "BOWLING GREEN" PLAN

Student assessment of teaching skills became mandatory at the College of Business Administration in 1969. This action followed faculty approval of the evaluation concept and definition of a rating scheme. (See Appendix A) In succeeding years, faculty have reaffirmed the principle of student evaluation of instructors and continued use of the original plan.

The evaluation instrument employed in the "Bowling Green" scheme is "open-ended" - that is, students are requested to describe their instructors' strengths or shortcomings, to identify approaches which might contribute to improved performance, and to assign a particular grade (A-4, B-3, C-2, D-1, or F-0). The grades assigned provide an index of performance. Open-ended evaluation forms (rather than structured, multiple response types of instruments) were adopted in order to permit students the opportunity to "tell it as it is" - to provide evaluators degrees of freedom to define for themselves what constitutes effective teaching; be it the personal qualities of the instructor or learning outcomes. Moreover, because the open-ended instrument allows students maximum freedom to describe what "turned them on (or off)," the form is thought especially helpful to faculty as guides to self-improvement.

Administration of the student evaluation scheme within the College of Business is a shared responsibility of the Dean's Office, Departmental Chairmen, faculty, and students. The Office of the Dean is charged with logistical aspects of the program which consist primarily of distributing and collecting evaluation instruments, recording ratings, developing College and Departmental performance indices, and forwarding forms to appropriate Department Chairmen. Departmental Chairmen and/or senior faculty, using evaluations as feedback instruments, counsel faculty about approaches to exploit strengths or to repair shortcomings. The system is designed to promote the counseling function, specifically as it relates to the attempts to assist junior faculty become more effective teachers. Faculty are responsible for assuring students that their participation in the evaluation process will not affect academic standings. To insure the integrity of the scheme, faculty are absent when students complete

evaluation instruments during the last week of classes, students do not identify themselves on forms, and faculty may review instruments only after final grades have been deposited with the Registrar. Students, for their part, are expected to participate in the program in a responsible and honest manner. In order to generate greater student participation in the evaluation process, the Dean's Advisory Council (a group consisting of the undergraduate leadership of honorary societies and professional organizations in the College) each quarter addresses letters to students prior to the distribution of evaluation instruments, requesting genuine cooperation.

II

HAS TEACHING EFFECTIVENESS IMPROVED?

If student evaluation contributes to an improvement in teaching effectiveness, presumably students would acknowledge improvements by assigning faculty higher ratings. Reflections of improvement should take the form of (at least) a gradual increase in the overall average value of ratings received by faculty. This is not to say that improvements will become evident from one period to the next. Faculty may be unable to react immediately to feedback, or the identity between effective teaching and the pay-off system may be unclear so that perceived inducements do not act as motivators. Existence of a lead-lag relationship between student evaluation and improved performance of (say) two periods, hence, would seem reasonable. Expectedly, the average College "grade" of 2.8 over time would tend toward 3.0 and beyond.

In order to develop an understanding of the effect of evaluations upon faculty instructional performance over time, regression equations using the

method of least squares were developed for two faculty groups. Moreover, using a standard t test we sought to determine whether regression coefficients were significant. Our hypothesis was that coefficients were equal to zero (Ho: B=0) - that is, no trend increments existed - by way of inference, teaching effectiveness had not improved. In applying these methodologies, we assumed that the criteria employed by students to judge teaching effectiveness remained unchanged and application of criteria over evaluation periods was uniform.

Group A consists of 50 teachers who were members of the College of Business Administration faculty at the time student evaluation was introduced, Winter Quarter 1969-1970, and have participated in the program all eight quarters. Instructional skills of this group were evaluated by 22,141 students who were members of 1,000 course offerings. Members of Group B are 13 faculty who joined the College in September 1970 and have taught during six consecutive quarters. Group B "grades" reflect the assessments of 4,317 students enrolled in 195 courses. Excluded from these groups are faculty who were not assigned instructional responsibilities for all quarters under consideration, who left Bowling Green after Winter Quarter 1969-1970, or who are Fall 1971 additions. By segmenting faculty according to the scheme described above, it was thought that the (potential) impact of student evaluation upon teaching effectiveness might be better observed. To illustrate - faculty who joined the College September 1970 received (for the most part) appointments at lower academic ranks, accepted the concept of a mandatory student evaluation system, and were committed to instruction of undergraduate programs.

Averages of student evaluation for Groups A and B are described below.

EXHIBIT I

MEAN VALUES[1]

(BY GROUP AND QUARTER)

Quarter	A	B
Winter 1969-1970	2.83	
Spring 1969-1970	2.90	
Fall 1970-1971	2.75	2.79
Winter 1970-1971	2.88	2.79
Spring 1970-1971	2.98	2.97
Fall 1971-1972	2.86	2.99
Winter 1971-1972	2.72	2.97
Spring 1971-1972	2.91	3.09
Total period	2.88	2.93

[1]Total "quality points" (A=4, B=3, C=2, D=1, F=0) for each course offering for each faculty member/total student evaluations.

Results of the test for significance appear in Exhibit II.

EXHIBIT II

TEST FOR SIGNIFICANCE[1]

Group	t	Critical Value (= .01	DF	Decision on Ho
A	1.3000	3.143	6	Accept
B	.8169	3.747	4	Accept

[1]See Appendices B and C for method and calculations.

On the basis of test results we accept the null hypothesis that regression coefficients of our regression equations are as low as 0. We conclude, therefore, that because the overall value of student ratings has not increased, student evaluation has not resulted in an improvement in teaching effectiveness.

III

IMPLICATIONS

The evaluation system employed at Bowling Green appears to have failed to stimulate an overall improvement in teaching. The question at hand then is -- what factors account for apparent failure of the teacher evaluation scheme to stimulate improved instruction? From our viewpoint, shortcomings in the overall administration of the student evaluation program and faculty capabilities and attitudes are factors which primarily account for the seeming inability of student evaluation to promote teaching effectiveness.

Before we discuss these factors, we need to re-examine our original assumption that student perceptions about faculty instructional performances are constant. The assumption may be invalid due to the possibility of changes in student expectations of teaching competencies. Students who have participated in the evaluation process for several quarters, for example, may now require more from their teachers and, therefore, might become more critical in their evaluation of performance. The possibility exists, then, that while teaching effectiveness may have improved it might not have been recorded as evaluators became harsher critics. Although we have reviewed thousands of student responses to questions asked on the evaluation form and, in turn,

have acquired some feel for student judgments about effective teaching, we do not believe sufficient information exists in order to affirm or deny the assumption of constant expectations.

The evaluation scheme assumes counseling relationships between faculty and Departmental Chairmen. Chairmen and/or "master" teachers are expected to assist faculty who receive low ratings to become more competent in the classroom. Although some Chairmen and faculty have worked diligently to help colleagues (especially junior members) become more effective teachers, indications are that neither counseling nor joint development of instructional skills is practiced widely. In some cases, little attention is directed to counseling activities because the concept of student evaluation of instructors is not accepted and, in other situations, efforts directed to the sharpening and honing of teaching skills is thought to be a high cost/low reward activity - in essence, some may perceive improved teaching efforts as performance punishing activities.

Good teaching is rewarded - in fact, recommendations for merit salary increases and promotions must be documented by results of student evaluation. Nevertheless, student evaluations have only recently served as inputs to decisions regarding definition of rewards (Group A - 1969-1972; Group B - 1970-1973). Chairmen and the College's administration could do more to demonstrate the relationship between teaching excellence and rewards. To illustrate: use of Management by Objectives as a joint (faculty-Chairman) goal determination and appraisal technique in place of the traditional scheme of faculty evaluation and allocation of rewards may be desirable to assist faculty develop and perceive

teaching improvement as performance non-punishing (rewarding) activities.

Results of student evaluation of instructors are not made public. Student assessments are seen only by the College's administration, the assessed faculty member, and his Department Chairman. One reason for restricting this information is that interpersonal comparisons may lead to conflict situations among faculty. But, of greater import, student evaluation is keyed to self-improvement.

The initiative to seek assistance from colleagues, as a means to self-improvement and self-development must come from the faculty member. In general, faculty are reluctant to seek the help of their colleagues. Senior faculty tend not to ask for help from junior faculty members who are recognized for their teaching ability because of status relationships associated with rank, age, and tenure differences.

Another attitudinal influence is that senior faculty are charged with the responsibility of evaluating overall performance of junior faculty. This relationship is perceived by some faculty (in particular, non-tenured members) to operate contrary to the development of mutual trust and confidence. In addition, the academic environment tends to stress individual contributions, a tendency which may frustrate attempts at team building and interdisciplinary efforts to promote greater teaching effectiveness.

With respect to the notion of faculty capabilities as a constraining influence upon the improvement of teaching effectiveness, we mean simply that some faculty do not possess the potential to markedly improve their performance. Students from one period to the next rate the collective College teaching effort as a "solid" B. We interpret this rating to mean that, as a whole, faculty are rendering fine contributions to the learning process - it

would be difficult to improve on this. Perhaps the performance discrepancy (desired-actual) is not significant enough to indicate existence of performance problems. First of all, we cannot expect faculty already achieving high ratings (say) 3.5 and above to enhance the value of their teaching. Further, we believe that it is difficult, via student ratings, for behavioral changes to occur in some faculty who, over time, were not effective teachers but, in terms of self-appraisal, perceived or rationalized that they helped students learn. Behavioral change occurs slowly and the eight quarter and six quarter evaluation periods for groups "A" and "B," respectively, may not be sufficiently long enough for faculty learning and experimentation results to be observed.

IV

CONCLUSION

The student evaluation system employed at Bowling Green appears not to be effective - although student evaluation of instructional skills may have helped some members of the faculty to become better teachers, from an aggregate point of view it does not appear that the evaluation scheme has resulted in improved teaching. Experiences with student evaluation at the College of Business should not, however, become arguments against the evaluation concept. The potential of student assessment of faculty instructional competencies to help faculty create a better environment for learning still appears to be significant. Realization of this potential, nevertheless, is dependent upon the improvement of Department Chairman-faculty counseling relationships, distribution of greater rewards for effective teaching, a change in attitude

of some faculty regarding students' abilities to define good teaching and describe approaches to improve performance, and faculty exploitation of the results of an evaluation system. Student evaluation does identify skill deficiencies. It appears the shortcomings of this scheme relate to faculty perceptions of the importance of teaching performance and administrative activities to help remove obstacles to creating an environment conducive to improving teaching effectiveness.

Karl E. Vogt and Harry Lasher
College of Business Administration
Bowling Green State University
Bowling Green, Ohio 43403

APPENDIX A

COLLEGE OF BUSINESS ADMINISTRATION
Bowling Green State University
Bowling Green, Ohio 43403

STUDENT EVALUATION OF INSTRUCTOR

COURSE NUMBER AND TITLE _____

INSTRUCTOR'S NAME _____

The faculty members of the College of Business Administration have voted to use student evaluations because they feel that through these evaluations teaching competence can be improved. Consequently, you are being asked to grade your instructor on his overall teaching effectiveness.

Indicate in a brief statement the factors which you consider significant about the instructor of this course.

How do you believe this instructor could improve his/her teaching effectiveness?

Please circle the letter grade which you assign to this instructor.

A	B	C	D	F
(Excellent)	(Good)	(Average)	(Poor)	(Unsatisfactory)

NOTE: THE RESULTS OF THIS RATING BY STUDENTS WILL NOT BE MADE KNOWN TO THE INSTRUCTOR UNTIL AFTER FINAL GRADES HAVE BEEN RECORDED.

APPENDIX B

Student Evaluation of Faculty Instructional Competencies Has Not Resulted In An Improvement In Teaching Effectiveness Over Time

(Group A)

MODEL $Y_i = \alpha + \beta X_i + \varepsilon_i$

TEST $H_o : \beta = 0$

$H_1 : \beta > 0$

$$b = \frac{N\Sigma XY - (\Sigma X)(\Sigma Y)}{N\Sigma X^2 - (\Sigma X)^2} \qquad a = \bar{Y} - b\bar{X}$$

X (Quarter in Transforming Units)	(Grade Point Average Of Student Evaluations)	
0	2.83	$\Sigma X = 28$
1	2.90	$\Sigma Y = 23.03$
2	2.75	$(\Sigma X)^2 = 784$
3	2.88	$\Sigma X^2 = 140$
4	2.98	$\Sigma Y^2 = 66.3303$
5	2.86	$\Sigma XY = 81.15$
6	2.92	
7	2.91	

$$b = \frac{8(81.15) - 28(23.03)}{8(140) - 784} = \frac{4.36}{336} = .0130$$

$$a = \frac{1}{8}\left[23.03 - .0130(28)\right] = \frac{22.6660}{8} = 2.8332$$

$$\therefore \hat{Y}_i = 2.8332 + .0130X_i$$

$i = 0$ At Winter Quarter, 1969-1970

$$t = \frac{b - \beta}{S_b}$$

$$S_b = \frac{Sy.x}{\sqrt{\Sigma(x_i - \bar{x})^2}} = \frac{\sqrt{\frac{\Sigma Y^2 - a\Sigma Y - b\Sigma XY}{n-2}}}{\sqrt{\Sigma(x_i - \bar{x})^2}} = \frac{\sqrt{\frac{66.3303 - 2.8332(23.03) - .0130(81.15)}{8-2}}}{\sqrt{\frac{336}{8}}} = \sqrt{\frac{.0044}{42}} =$$

$$\sqrt{.0001} = .01$$

$$t = \frac{.0130 - 0}{.01} = 1.30$$

$t_{.99(6)} = 3.143$

DECISION: ACCEPT H_o (REGRESSION NOT SIGNIFICANT)

APPENDIX C

Student Evaluation of Faculty Instructional Competencies Has Not Resulted In An Improvement In Teaching Effectiveness Over Time

(Group B)

MODEL $Y_i = \alpha + BX_i + \epsilon_i$

TEST $H_o : \beta = 0$

$H_2 : \beta > 0$

$$b = \frac{N\Sigma XY-(\Sigma X)(\Sigma Y)}{N\Sigma X^2 - (\Sigma X)^2} \qquad a = \bar{Y}-b\bar{X}$$

X (Quarter in Transforming Units)	Y (Grade Point Average of Student Evaluations)	
0	2.79	$\Sigma X = 15$
1	2.79	$\Sigma Y = 17.62$
2	2.97	$(\Sigma X)^2 = 225$
3	2.99	$\Sigma X^2 = 55$
4	2.99	$\Sigma Y^2 = 51.8174$
5	3.09	$\Sigma XY = 45.11$

$$b = \frac{6(45.11)-(15)(17.62)}{6(55) - 225} = \frac{5.36}{105} = .0510$$

$$a = \frac{1}{6}\left[17.62 - .0510(15)\right] = \frac{17.5435}{6} = 2.9239$$

$$\therefore \hat{Y}_i = 2.9239 + .0510X_i$$

i = 0 At Fall Quarter, 1970-1971

$$t = \frac{b-\beta}{S_b}$$

$$S_b = \frac{Sy.x}{\sqrt{\Sigma(x_i-\bar{x})^2}} = \frac{\sqrt{\frac{\Sigma Y^2-a\Sigma Y-b\Sigma XY}{n-2}}}{\sqrt{\Sigma(x_i-\bar{x})^2}} = \frac{\sqrt{\frac{51.8174-2.9239(17.62)-.0510(45.11)}{4}}}{\sqrt{\frac{105}{6}}} =$$

$$\sqrt{\frac{.0682}{17.5}} = \sqrt{.003897} = .06243$$

$$t = \frac{.0510-0}{.06243} = .8169$$

$t\ 99(4) = 3.747$

DECISION: ACCEPT H_o (REGRESSION IS NOT SIGNIFICANT)