

## Cornell University School of Hotel Administration The Scholarly Commons

---

Articles and Chapters

School of Hotel Administration Collection

---

1988

# Improving the Teaching Evaluation Process: A Report from the Classroom

Cathy A. Enz

*Cornell University School of Hotel Administration, cae4@cornell.edu*

Suzyn Ornstein

*Suffolk University*

Barbara Allen

*Carleton University*

Follow this and additional works at: <http://scholarship.sha.cornell.edu/articles>

---

### Recommended Citation

Enz, C. A., Ornstein, S., & Allen, B. (1988). *Improving the teaching evaluation process: A report from the classroom*[Electronic version]. Retrieved [insert date] from Cornell University, School of Hotel Administration site: <http://scholarship.sha.cornell.edu/articles/625>

This Article or Chapter is brought to you for free and open access by the School of Hotel Administration Collection at The Scholarly Commons. It has been accepted for inclusion in Articles and Chapters by an authorized administrator of The Scholarly Commons. For more information, please contact [hlmdigital@cornell.edu](mailto:hlmdigital@cornell.edu).

---

# Improving the Teaching Evaluation Process: A Report from the Classroom

## **Abstract**

[Excerpt] Students' evaluations of instructors' teaching may be one of many longstanding approaches to improving instruction which has lost its utility as a diagnostic tool and has become a quantitative device used to determine merit pay and tenure. Comments such as "Instructor has great legs" or "You should lose weight" or "Buy a new tie," while occasionally amusing, suggest that the students themselves take the process of feedback lightly. Instructors who view student responses as critical to their future salary and promotion opportunities may take a tactical stance and pander to student wishes: if the material is too complex, make it simpler; if the assignments are too long, make them shorter; if the grade distribution is too discriminating, adjust it. Perhaps more distressing is the good teacher who will not tinker or innovate for fear that changing the course will result in lower student evaluations. Across the country the process of evaluating instruction has moved from an activity centered on providing helpful feedback and suggestions for improvement into a means of making comparisons across faculty in an effort to link salary and promotion decisions to seemingly objective data.

## **Keywords**

teacher evaluations, student feedback, business school

## **Comments**

### **Required Publisher Statement**

© SAGE. Final version published as: Enz, C. A., Ornstein, S., & Allen, B. (1988). Improving the teaching evaluation process: A report from the classroom. *Journal of Management Education*, 12(4), 114-123.  
doi:10.1177/105256298801200415

Reprinted with permission. All rights reserved.

# IMPROVING THE TEACHING EVALUATION PROCESS: A REPORT FROM THE CLASSROOM

Cathy a. Enz

*Indiana University*

Suzyn Ornstein

*Suffolk University*

Barbara Allen

*Carleton College*

...the very activity [teaching] brings with it the danger of preferring teaching to knowing, of adapting oneself to what the students can or want to learn, of knowing oneself only by one's students. (Bloom, 1987)

Students' evaluations of instructors' teaching may be one of many longstanding approaches to improving instruction which has lost its utility as a diagnostic tool and has become a quantitative device used to determine merit pay and tenure. Comments such as "Instructor has great legs" or "You should lose weight" or "Buy a new tie," while occasionally amusing, suggest that the students themselves take the process of feedback lightly. Instructors who view student responses as critical to their future salary and promotion opportunities may take a tactical stance and pander to student wishes: if the material is too complex, make it simpler; if the assignments are too long, make them shorter; if the grade distribution is too discriminating, adjust it. Perhaps more distressing is the good teacher who will not tinker or innovate for fear that changing the course will result in lower student evaluations. Across the country the process of evaluating instruction has moved from an activity centered on providing helpful feedback and suggestions for improvement into a means of making comparisons across faculty in an effort to link salary and promotion decisions to seemingly objective data.

If, as we have suggested, teaching evaluations are used for evaluation versus feedback, students do not seriously complete the forms, and instructors utilize data in order to water down and develop courses that are popular but not educational — should we then rid ourselves of this perverted instrument of teacher assessment? We believe the answer is clearly no, but it is essential that the purpose and usefulness of the teaching evaluation process be examined and suggestions for change be contemplated.

In this paper we will discuss the purpose of teaching evaluations, provide the findings of a recent study examining the perceptions of organizational behavior faculty concerning their

institutional approaches to evaluation, and conclude with some suggestions for improving three aspects of evaluation, including the design of evaluation forms, new formats for feedback, and strategies for learning from evaluations.

Most business schools across the country require some evaluation of faculty teaching by students. These student evaluations serve numerous purposes, but are used primarily to provide superiors with some indication of teaching skills and the individual professor with feedback. Student opinions are valued because (1) they provide a seemingly quantifiable indicator of teaching excellence to be used to document promotion and tenure materials, (2) few alternative methods exist to obtain responses from a large number of persons exposed to a particular faculty member, and (3) they offer a means of confirmation and diagnosis to the instructor.

### **What do Teaching Evaluations Measure?**

Research into the development of teaching evaluations has revealed that instruments which evaluate the instructor do not provide for data on improving instruction (Aubrecht, 1979). Improving instruction is not usually the purpose behind rating scales designed to evaluate current teacher performance. Many forms are not written so as to be simultaneously evaluative and diagnostic. In fact, the idea that these purposes may differ and require different instruments is often not considered. Given that evaluation is often preferred over diagnosis and that most business schools rely on evaluative forms to facilitate collection of data for salary decisions, little formal attention is given to providing the instructor with suggestions for improvement and feedback on specific behaviors, and salary may not be linked to improvement in teaching. Moreover, as an evaluative tool, student preferences may themselves be a limited indicator of teaching excellence.

According to researchers (Feldman, 1978; Wollin & Montagne, 1981), extraneous factors such as class size, academic field, and the physical environment in the classroom also affect student ratings. Smaller classes and liberal arts courses are associated with higher ratings. In addition, classrooms that are brightly painted and decorated with posters, rugs, cushions, and plants result in significantly higher teaching evaluations than bland classrooms.

The tension between what appears practical, useful, and interesting in the view of the student and the teacher's vision of the necessary aspects of the discipline's canon of knowledge which must be taught is also expected to influence the evaluation process. Different departments within business schools may yield higher or lower ratings as a result of the subject matter and the fit between the discipline's body of knowledge and student expectations. In behavioral courses, for instance, students often see the duplication of what they have learned in psychology. In such cases, the instructor is obviously not the only source of knowledge to the student. Thus, comparative and baseline ratings between different faculty members in diverse areas may prove misleading.

An instructor's approach to teaching can also play an unmeasured role in the interpretation of evaluations. Do instructors who rely on cases and discussion have problems when compared to professors who lecture? We believe the answer is yes when an evaluation question asks whether "This instructor paces his/her lectures effectively." At Indiana University, where this question is used, the category "undecided" represents a neutral position. Since there is no response category of "not applicable" for this item, a mean score near 4 on a 7-point scale could mean that either the question was irrelevant or the instructor was mediocre in the pacing of lectures. When interclass comparisons are made, this type of item becomes difficult to interpret. These problems also exist when questions presume that exams were given and textbooks used. The conventional nature of these questions penalizes the instructor for innovating.

What do teaching evaluations measure anyway? From our discussion so far it is evident that they cannot provide evaluative and diagnostic material at the same time. In addition, teaching evaluation forms do not control for extraneous factors and pedagogical variety. While we have highlighted a few of the more obvious limitations of teaching evaluations, they are still used as a primary — and in some cases sole — means of determining salary distributions. Clearly evaluations are important, albeit flawed, devices for assessing teaching activity.

### **A Study of Teaching Evaluation**

To explore the nature of and satisfaction with teaching evaluations in business schools across the country, we conducted a study of organizational behavior professors. A 46-item questionnaire was developed to capture the attitudes of OB professors toward the teaching evaluation process. The survey was designed to obtain information on the following: (1) how the instructor's course is designed, (2) the nature of and level of satisfaction with institutional teaching evaluations, and (3) reactions to a variety of student comments.

A total of 127 college professors responded to the survey drawn from a list of 215 participants in the 1986 Organizational Behavior Teaching Conference (OBTC), and ten OB instructors from Indiana University and Southern Methodist University. More than two-thirds of the sample were males averaging 43 years of age and twelve years of teaching experience at the college level. Respondents taught in both public (54.4 percent) and private (45.6 percent) institutions with average student bodies of 15,000 and class sizes of around 43 students per class. Most instructors taught both upper-class undergraduates and graduate students.

#### **Design of the Course**

A large percentage of respondents (84.7 percent) teach courses primarily for business students and rely on a discussion approach augmented by other activities. Lectures and case analyses were second and third in popularity as pedagogical tools for instruction. While many different approaches exist for testing student performance, the most popular strategy was the use of multiple choice and true/false exams. Essay exams and term papers were also popular means of evaluating students. Games and simulations were the least popular approaches to

student evaluation. In sum, the typical respondent teaches business students relying primarily on a discussion approach and evaluating students with multiple-choice exam questions.

### Attitudes Toward Teaching Evaluations

Almost all (99.2 percent) of the respondents have their students complete teacher evaluations, which are usually distributed once during a term and use machine-coded and scored items. These evaluations are often (82.3 percent) used to assess teaching performance as opposed to providing feedback simply for the instructor's improvement. Most institutions appear to rely on teaching excellence information for purposes of salary decisions; hence, student responses to these computerized evaluation forms are critical in determining teaching-based institutional rewards which the instructor receives on a yearly basis.

Almost two-thirds (65.8 percent) of the respondents were very or somewhat satisfied with the evaluation measure used in their university, while 34.2 percent were somewhat or very dissatisfied. Interestingly, the sample was evenly split on whether the evaluation method used told them what they needed or wanted to know. A total of 56.4 percent felt the evaluations gave them adequate information on teaching, while 42.7 percent felt they did not. The evaluations were felt to represent accurately the teaching of over 75.9 percent of those surveyed. In sum, respondents are reasonably satisfied with their teaching evaluations and regard them as accurate, but many feel that the current evaluations do not provide sufficient information to assess their teaching. In particular, respondents indicated that they would like to have suggestions on how to improve their teaching, more detail about specific behaviors and activities used in the course, suggestions for what to retain, change and delete, and assessments of the student's opinion about her or his individual learning.

Relationships between demographic characteristics of teachers and their institutions and satisfaction with current teaching evaluations and the accuracy of these evaluations were examined using one-way Analyses of Variance (ANOVA's). The findings indicate that there are no significant differences in satisfaction or accuracy on the basis of gender, class size, school mission (teaching versus research), or teaching approach (discussion, lecture, or case analysis). Hence, male and female instructors, long-time and new teachers, those with small versus large class sizes, and those who rely on lecture as opposed to discussion and case analysis do not differ in their levels of satisfaction with evaluations or their perceptions of the accuracy of these evaluations.

Significant differences in satisfaction ( $F = 39.53, p < .001$ ) and accuracy ( $F = 34.35, p < .001$ ) were found between respondents who wanted more information from the evaluation form and those who did not. Respondents who felt that the teaching evaluations should provide more information were less satisfied with the current evaluation measure and questioned the evaluation form's accuracy.

### Reactions to Student Comments

Using a list of twenty common student comments, each respondent was asked to indicate whether they take action, ignore, or do not receive the listed comments. The frequency of responses in each category is presented in Table 1. The typical comments for which professors take action are: (1) the evaluation system was ambiguous, (2) the textbook was terrible, (3) I feel I didn't learn very much, (4) topics covered were not integrated or organized, (5) feedback about grades were poor, and (6) instructor allowed a few students to

Comment	Frequency		
	Take Action	Ignore	Don't Receive
Subject matter is common sense	23	35	63
Ambiguous evaluation system	54	11	57
Boring subject	30	21	73
Textbook was terrible	80	6	39
Instructor was not interested in students	32	4	88
Course required more time than I thought	23	72	29
I feel I didn't learn very much	43	23	58
Instructor was not available outside of class	39	19	65
Feedback about grades were poor	42	6	75
Examinations didn't cover lecture material	29	12	83
Instructor didn't know much about the subject	12	6	105
Instructor was not prepared for class	18	4	102
Topics covered were not integrated or organized	49	7	68
Instructor talked too fast (slow)	40	9	75
Instructor didn't allow for discussion or disagreement	21	4	89
Instructor took too long to return papers or exams	32	9	83
Too much focus on one aspect of the subject	20	10	94
There were not enough (too many) cases, exams, papers	13	29	57
I thought this class would be about _____ and it wasn't	13	25	64
Instructor allowed a few students to dominate group discussion	49	11	64

**Table 1** Reaction to Student Comments

dominate group discussion. The most frequently ignored comments were: (1) course required more time than I thought, (2) subject matter is common-sense, and (3) there were not enough (too many) cases, exams, papers. Finally, the comments which respondents usually did not receive were: (1) instructor didn't know much about the subject, (2) instructor was not

prepared for class, (3) instructor didn't allow for discussion or disagreement, and (4) too much focus on one aspect of the subject.

A total of 77.5 percent of those surveyed have received irrelevant or negative comments from students. Examples cited include "instructor is unfair — he made me think," "class was a waste of time," "too women's lib," and "why did this cheapskate school hire a pinko, commie, bum?" When asked how they deal with these comments most respondents suggested they ignore negative comments. Others offered the following coping strategies: laughter, tears, focus on the positive, get defensive, talk with colleagues, rationalize, discuss "cheap shots" with students prior to evaluations, and work on better impression management.

In sum, the survey revealed that OB instructors at the end of every term use a machine-scored evaluation form which they are generally satisfied with, but which fails to yield helpful suggestions for improving the course or enhancing instruction. Those instructors most troubled by their current evaluation forms are faculty who question the accuracy and usefulness of the student comments and lament the absence of feedback in the process. Given the survey respondents' concerns over the lack of diagnostic prescriptions to improve teaching, in the remainder of this paper we will explore alternatives and modifications to the teaching evaluation forms so popular across the country.

### **Suggestions for Change**

There are numerous possibilities for improving teaching evaluations. However, any suggestions for change must take into account that the vast majority of institutions use teaching evaluations specifically for the purpose of performance appraisal. Thus, it is necessary to differentiate between the types of changes that will be practical for improving the evaluation process and those changes that will facilitate the feedback process. Based on this caveat, suggestions for improving evaluations can be divided into three categories: improving evaluation forms and their implementation, developing new forms and formats for feedback focused on the development of teaching skills, and introducing strategies for learning from evaluations.

#### **Improving Evaluations and their Implementation**

The improvement of standardized, computer-scored evaluation forms can be accomplished in various ways. Questions should be pilot-tested to ensure that they are clear and unambiguous to students. Questions also must be relevant to the characteristics that differentiate excellent from poor instruction. The best advice from performance appraisal experts is that questions should be behavioral in nature. Of course, all scaled types of evaluation surveys should be examined for reliability and validity. While this appears to be obvious, few instruments measuring teaching performance have excellent psychometric properties. A simple way of improving the reliability of these evaluation forms is to assure that more than one question is provided to capture any specific construct. There should also be agreement by those developing the instrument regarding what specific aspects of excellent



teaching are measured by the questions; that is, what are the constructs being tapped. Finally, due to the great variability in pedagogical tools, if questions are asked about specific teaching methods (e.g., lecture) there must be a space to respond “not applicable” so that mediocre scores can be differentiated from those signifying the question’s irrelevance.

Some steps should be taken to help increase student awareness of the importance of their seriously completing teaching evaluations. One such step is to have students actually sign their evaluations. This can be done while maintaining anonymity by having forms that are numbered, and having identical numbers in a tear-off sheet attached to the form. This process suggests to students that their input is being taken seriously. It also provides a faculty review committee with valuable information if, for instance, there are students who have developed reputations as agitators or naysayers. One dilemma which might be difficult to resolve in this endeavor is to distinguish between legitimate dissenters and troublemakers. The final benefit of this method is that it provides evaluators the opportunity to discuss specific evaluations with students and to ask clarifying questions.

Another suggestion for the improved implementation of evaluations is for these evaluations to be administered by other faculty members (or better yet, the dean of academic affairs!). Students devalue the import of evaluations when they are haphazardly administered by other students or even secretaries. An independent third party also frees the faculty member from the uncomfortable administrative task of distributing and collecting evaluations. By giving this responsibility to a neutral party students may be more willing to be honest. The flagrant misuse of teaching evaluations by faculty who discard negative responses or retaliate against students who express criticism is also eliminated when a third party handles the evaluation activity.

### *New Forms and Formats for Feedback*

Although it is possible to envision an improved evaluation form, it is still that — a form for evaluation. Some of our respondents suggested individualized methods for obtaining specific teaching-skill related feedback such as small group discussions by students with a written summary provided to the teacher. Because many of the respondents to our survey suggested that they would also desire teaching-specific feedback, it is necessary to provide suggestions for new forms and formats of feedback. The simplest and most straightforward suggestion is to design your own feedback form. One of the authors does this quite successfully by asking students to rate class assignments, exercises, cases, exams, and readings on a scale indicating their value to the students (“extremely valuable” to “no value whatsoever”). In addition, students rate each item as to whether or not it should be repeated in future classes (“definitely should be repeated” to “definitely should not be repeated”). Based on this, future decisions about course design are made.

One respondent to our questionnaire suggested that he made a weekly student evaluation of the course activities part of his routine. He asked for the students to comment on

straightforward questions such as: “What did I learn this week?” or “How will \_\_\_\_\_ topic be useful to me?” or “What was valuable/valueless in what we did this week in class?” The instructor may still be in a quandary about how to reconcile the students’ views of utility with her or his own ideas of knowledge, but the teacher will at least know what remains a cloud of abstraction and what has been taken to heart by the students.

Because the evaluation of teaching is an attempt to quantify the intangible, it is unrealistic to expect Likert-type responses to general questions to provide much information that is truly informative. One of the authors uses open-ended questions to allow for more class-specific and student-guided feedback. A simple form asks students to indicate what they liked most and least about the class and instructor. Another question focuses on suggestions for change by asking, “If you could change one thing in the class what would it be?”

Both of the suggestions for feedback-guided forms give the faculty member an opportunity not only to receive feedback but also to respond to it. This process can be made even more effective by collecting feedback midway, or even more frequently, through the term. The instructor who is capable of making adjustments to his/her teaching approach and material serves as a role model to the students who are expected to change when they receive evaluation feedback during the term. A final result of utilizing forms oriented to teaching-skills or specifics of course design is that students feel more positive about the instructor and the class may run more smoothly, leading to a better learning environment for all.

Other than paper and pencil forms, there are other feedback formats available. Peer evaluation, although generally unpopular from the perspective of faculty, may be much more acceptable if the purpose is feedback and growth. Of course, the success of this depends on one’s respect for the “peer’s” teaching. One way to maximize the benefits of peer evaluation is to invite a colleague to a class session in which you are presenting new and different material in which he/she may have an interest. Educational specialists in other parts of the university may also prove valuable evaluators of the process, if not the content, of one’s teaching. Another format for generating feedback is through class discussion. Students may meet with or without the instructor, as a class or in small groups. The purpose of these meetings is to develop a feedback report—that is, information that the instructor can use to improve his or her teaching. One respondent to our study suggested group discussion of the evaluation questions prior to the student’s individual completion of the form. He believed the students’ views were more accurate by deliberation through the process of airing complaints and commendations in a discussion format with peers.

### *Strategies for Learning from Evaluations*

Although in our sample relatively small percentages of instructors indicated that they ignored certain student comments, it is to be expected that less dedicated (and perhaps more honest) colleagues do ignore student evaluations. This suggests that perhaps evaluations for both feedback and evaluative purposes are not taken as seriously by us as they should be. As

was the case with other problems of evaluation, this problem also has various remedies. One solution is to use the evaluation as a teaching tool. This can be done by using the evaluation format to teach about performance appraisal. As part of an exercise in job analysis and evaluation, students could devise a measure of the job of teacher. This is a reasonable exercise, as most of our business school students will be employed in occupations where evaluating subordinates' performance will be based on less than objective, quantifiable standards. Incorporating the evaluation into an exercise of this type teaches the students about performance appraisal and may result in much more meaningful feedback to the instructor about what students think is good teaching. Another possibility is to use the evaluation process as an opportunity to study goal setting. At the beginning of the term the students and instructor could develop mutual goals for the term. The evaluation could then focus on the achievement of these goals.

Outside of the classroom there are also opportunities to benefit from evaluations. A "brown bag" series in which faculty discuss teaching could be useful. Additionally, there are two publications that may help to enhance your growth potential from evaluations: Kansas State University's Center for Faculty Evaluation and Development has published IDEA papers that cover topics related to the improvement of teaching and evaluation, and "The Teaching Professor" is a monthly newsletter that addresses issues relevant to teaching and its evaluation.<sup>1</sup>

### **Conclusion**

In summary, there are several approaches that we as teachers can use to improve the quality of our teaching evaluation system. Of course, all of these methods require a proactive stance on our parts. The great benefits to our placing time and energy in improving our evaluation systems is in the vastly improved feedback we may receive and the consequent improvement of our teaching and therefore our students' learning.

### **Notes**

<sup>1</sup>The IDEA papers can be obtained from the Center for Faculty Evaluation and Development at Kansas State University. The "Teaching Professor" is a newsletter published twelve times a year by Magna Publications, Inc., 2718 Dryden Drive, Madison WI., 53704-3006.

### **References**

- Aubrecht, J.D. (1981). Reliability, validity and generalizability of student ratings of instruction. *IDEA Paper #6*. Kansas State University Center for Continuing Education: Center for Faculty Evaluation and Development.
- Bloom, A. (1987). *The closing of the American mind*. New York: Simon & Schuster.
- Feldman, K.A. (1978). Course characteristics and college students' rating their teachers and courses: A review and analysis. *Research in Higher Education*, 9, 199-242.

Wollin, D.D., & Montagne, M. (1981). College classroom environment: Effects of sterility versus amiability on student and teacher performance. *Environment and Behavior*, 13, 707-716.