AC 2011-334: REEVALUATING THE STUDENT EVALUATION PROCESS

Scott Dunning, University of Maine

Dr. Dunning is the Director of the University of Maine School of Engineering Technology (SET). He is responsible for overall supervision of the four academic programs within the school and serves as its academic dean. He is also a tenured Professor of Electrical Engineering Technology. His primary academic responsibilities have been for teaching courses within the electrical engineering technology program.

Dr. Dunning's primary research interests are in the areas of power systems optimization and the application of energy efficient technologies to industry. Previously, he was the Director and Founder of the University of Maine Industrial Assessment Center where he supervised the completion of over 200 industrial assessments for New England manufacturers. In 2002, the center evolved into the University of Maine Advanced Manufacturing Center. As Executive Director of the AMC, he coordinated engineering outreach projects to support economic development in the state as well as supporting research centers on the University of Maine campus.

Dr. Dunning is a registered Professional Engineer in the State of Maine and previous Chair of the Maine Section of the Institute of Electrical and Electronics Engineers. He is a Certified Energy Manager and provides C.E.M. training for the Association of Energy Engineers both nationally and internationally. He is also previous National Chair for the Engineering Technology Division of the American Society for Engineering Education. He currently serves as a commissioner for the Technology Accreditation Commission of ABET, inc.

Reevaluating the Student Evaluation Process

Abstract

University faculty members are typically evaluated through their contributions to the teaching, scholarly research and public service activities of their institution. Expectations vary by institution for numbers and types of scholarly publications, proposals submitted and funded, internal and external service, etc. However, student evaluations of teaching are almost universally accepted as a performance measure of teaching quality. The first rating scales for evaluating teachers were developed by Herman Remmers in 1927. [1] In the eighty years following this first development, the SET process has been refined and studied extensively. This paper discusses the benefits of a short form with objective questions that requires supporting comments for positive or negative ratings.

Issues with Student Evaluation Forms

Student evaluations of teaching are typically used for two purposes. The first purpose is to provide feedback to individual faculty members from the student audience. The second purpose is for department peers and administrative supervisors to evaluate the quality of teaching provided by a faculty member. The usefulness of student evaluations to objectively meet these dual purposes has often been criticized. The most common criticism is that students tend to provide higher scores when they expect a higher grade in the course. This link has proven to be valid by various studies.[2,3]

This has led to the criticism that student evaluations of teaching are just a popularity contest and if a faculty member wishes to improve their ratings, they need to grade easier. As students receive higher grades, they feel greater praise which causes them to like the instructor more. Some studies have shown this criticism to be valid.[4] One study revealed that 70% of students in one study admitted their expected course grade affected their perception of the instructor. [5] This raises significant concerns that faculty may remove difficult course content to improve their ratings. One faculty survey provided data demonstrating that this occurs.[6]

The good news is that weakening course content is not the only factor that can affect student ratings. Student ratings appear to also be significantly affected by the presentation style of the instructor. William and Ceci performed an experiment that demonstrated how presentation style changes can positively affect student evaluations. [7] Variation in voice inflection and use of nonverbal gestures was demonstrated to increase student engagement and improve student perception of an instructor's teaching effectiveness.

Developing Effective Evaluation Forms

For this paper, data received using a standard form approved by the institution was compared to data received from a non-standard evaluation form. The standard institutional evaluation form consists of a one page form with 29 questions. At the end of the form, a box measuring 2" by 4" is provided for input.

The standard form was developed to provide questions designed by faculty that target the most likely areas of student comment. The results are in the form of mostly numeric data which is easily compiled by the administration. The data can then be examined and conclusions drawn to identify areas for improvement and areas of excellence. Faculty scores can be compared against prior scores or against peer ratings.

The standard form has several weaknesses identified by faculty. First, the excessive number of questions may cause students to lose interest when filling out the form leading to inaccurate results. Second, if students do not read the questions carefully, they may answer the question inaccurately. For half of the questions on the form, a score of 1 is a positive indicator. For the other half of the questions, a 5 is a positive indicator. Third, room is not provided for students to write comments to justify each score. Thus, mistaken answers are not easily identified. Finally, instructor results are evaluated by average results. Those unfamiliar with statistics may misinterpret the data and draw false conclusions.

An alternative form was developed to address the identified weaknesses. In the nonstandard form, there are only five questions. Each question has a numeric score for selection from (1)-outstanding to (5)-unsatisfactory. Additionally, there is a large area for comments provided after each question. To ensure accuracy, scores above or below (3)-average are required to be justified by a comment or they are not scored.

To compare the value of data provided by each form, students in one course were asked to fill out both forms during the spring semester of 2010. Table 1 presents the results from the standard form. Questions 2, 4, 12, 13, and 29 are highlighted since they correspond closely to questions 1 through 5 on the nonstandard form. Table 2 presents the results from the nonstandard form. Table 3 presents a comparison of the scores for the selected questions. The numerical results show slight differences that may be due to slight variations in the wording of the questions.

	Traditional Student Evaluation (29 Questions)								
#	Description	1	2	3	4	5	Unrated	Average	Note
1	How prepared was the instructor for class?	27	4	1				1.19	1 good
2	How clearly were the objectives of the course presented?			2	10	19	1	4.41	5 good
3	How enthusiastic was the instructor about the subject?			6	11	14	1	4.13	5 good
4	How clearly did the instructor present ideas and theories	19	10	3				1.50	1 good
5	How much were students encouraged to think for themselves?	10	7	12	2		1	2.13	1 good
6	How concerned was the instructor for the quality of his or her teaching?			4	15	14		4.44	5 good
7	How orderly and logical were the instructor's presentations of the material?			3	7	21	1	4.44	5 good
8	How open was the instructor to other viewpoints?			8	10	14		4.19	5 good
9	Did the instructor show respect for the questions and opinions of the students?	21	9	1	1			1.44	1 good
10	How often were examples used in class?			2	3	27		4.78	5 good
11	Did the instructor inspire confidence in his or her knowledge of the subject?			2	5	25		4.72	5 good
12	How genuinely concerned was the instructor with students' progress?	12	12	7	1			1.91	1 good
13	Overall, how would you rate the instructor?	21	9	1	1			1.44	1 good
14	Were class meetings profitable and worth attending?			4	6	22		4.56	5 good
15	How would you rate the subject matter of this course	13	9	9	1			1.94	1 good
16	Did you develop significant skills in the field as a result of taking this course?	19	9	3	1			1.56	1 good
17	How was the pace at which the materials in the course were covered?			25	7			3.22	3 good
18	What is our overall rating of the primary textbook?	2		15	10	5		3.50	5 good
19	Were students required to apply concepts to demonstrate understanding?	19	12	1				1.44	1 good
20	How did the work load for this course compare to that of others of equal credit?		1	22	8	1		3.28	3 good
21	How much intellectual discipline was required in this course?		1	6	18	7		3.97	5 good
22	What is your overall rating of this course?	11	13	7			1	1.81	1 good
23	How promptly were assignments and tests returned?	9	6	13	3		1	2.25	1 good
24	Could tests be completed in the allotted time?	1	1	1	2	26	1	4.50	5 good
25	Did the instructor let you know what he or she expected on tests and assignments?	21	6	3	1		1	1.44	1 good
26	Did exams reflect the important aspects of the course?			2	8	21	1	4.47	5 good
27	How clear were examination questions?			2	7	22	1	4.50	5 good
28	How fair were the grading procedures?	19	7	4	1		1	1.53	1 good
29	Overall, how would you rate the examination procedure?	18	8	5			1	1.53	1 good

Table 1 – Results from the Standard Form

#	Description	1	2	3	4	5	Unrated	Average	Note
1	Class lectures or class activities were related to the objectives of the course.	10	0	5	1	0	1	1.81	1 good
2	The presentation of course material was clear.	5	4	8	0	0	2	2.18	1 good
3	Available time was used productively.	4	2	8	1	0	3	2.40	1 good
4	Assignments were challenging, relevant, and helpful.	3	6	5	0	0	4	2.14	1 good
5	All things considered, this instructor was effective.	4	6	5	0	0	3	2.07	1 good

Table 2 – Results from the Nonstandard Form

#	Description	1	2	3	4	5	Unrated	Average	Note	Long Form #	Old Form Avg.	Note
1	Class lectures or class activities were related to the objectives of the course.	10	0	5	1	0	1	1.81	1 good	2	4.41	5 good
2	The presentation of course material was clear.	5	4	8	0	0	2	2.18	1 good	4	1.5	1 good
3	Available time was used productively.	4	2	8	1	0	3	2.40	1 good	13	1.44	1 good
4	Assignments were challenging, relevant, and helpful.	3	6	5	0	0	4	2.14	1 good	29	1.53	1 good
5	All things considered, this instructor was effective.	4	6	5	0	0	3	2.07	1 good	12	1.91	1 good

Table 3 – Results Comparison

The major difference seen between the results for the two forms comes from the comments received. As mentioned previously, the standard form provides one small box for students to provide comments. This does not make for easy correlation with individual questions. It also discourages comments by providing little room. A summary of the comments received is provided in Table 4. While the four comments received were flattering, they provide little useful input for continuous improvement.

Professor X is extremely intelligent and conveys his knowledge very well.
Best teacher I had this semester.
Professor X is excellent and very well versed in the material and knows how to get it across
Professor X is an excellent professor, I learned a lot from this course. The choice of text was very clearly wrote. I would take any class with this man teaching

Table 4 – Comments from Standard Form

The nonstandard form provided far more comments for consideration. Tables 5 through 9 illustrate the range of comments provided for each of the five questions on the form. As can be seen, the comments provide useful information identifying the positive aspects of the instruction provided by the instructor. In addition, there are two additional questions on the form that do not include numerical scores but provide an open format for comment.

Class was always worth attending
Homework was very much like exam questions
Course objectives were always talked about while doing examples in class, homework, and on tests. Clear statements as to what is expected of us upon completing this course.

Table 5 – Question 1: Class lectures or class activities were related to the objectives of the course.

Good Professor
Some things take time and practice to understand
Instructor made sure all questions were answered before moving on to new material
The way the instructor explained the material covered in this course was very detailed and very easy to follow and understand. I was very impressed with his method of teaching.
Always provided example packets

Table 6 – Question 2: The presentation of course material was clear.

There was no fooling around.
Never any down or idle time, class was always filled with examples of current material

Table 7 – Question 3: Available time was used productively.

Very effective.
If you could do the homework, you could do the test.

Table 8 – Question 4: Assignments were challenging, relevant, and helpful.

Very effective. I have learned a lot from this course.
Great teacher.

Table 9 – Question 5: All things considered this instructor was effective.

Discussion

Before discussing the results, the limitations of the present study are acknowledged. First, the data provided is for a relatively small sample when compared to results compiled for the whole institution. However, there is no reason to believe that the results would be dissimilar to those provided from an expansion of the study. Second, there is a concern that by forcing students to

justify scores other than three they may be inclined to just answer three to reduce their effort. This question will be examined closer in a study currently under review.

The data in the current study suggests that the nonstandard form may provide significant improvement over the standard form. By asking a reduced set of questions while requiring comments supporting positive or negative scores, the nonstandard form appears to provide more useful information for continuous teaching improvement.

Bibliography

1.   Remmers, H.H., (1927), "The Purdue Rating Scale for Instructors", *Educational Administration and Supervision,* 6*,* 399-406.
2.   Wilson, R., (1998), "New Research Casts Doubt on Value of Student Evaluations of Professors", *Chronicle of Higher Education,* A12.
3.   Rice, L., (1988), "Student Evaluation of Teaching: Problems and Prospects", *Teaching Philosophy,* 11, 329-44.
4.   D'Apollonia, S., Abrami, P.C.,(1997), "Navigating Student Ratings of Instruction", *American Psychologist*, 52, 1198-1208.
5.   Gilbaugh, J.W., (Feb. 1982), "Renner Substantiated", *Phi Delta Kappan*, 63, 428.
6.   Ryan, J.J., Anderson, J.J., Birchler, A.B., (December, 1980), "Student Evaluations: The Faculty Responds", *Research in Higher Education*, 12, 317-33.
7.   Williams, W.M., Ceci, S.J., (Sept./Oct. 1997), "How'm I Doing? Problems with Student Ratings of Instructors and Courses", *Change: The Magazine of Higher Learning*, 29, 12-23.