The Effective Teaching Committee of the Faculty Senate conducted a survey of all full- and part-time instructional faculty in May/June 2014. Our charge was to obtain feedback on the usefulness of the Course Evaluation form for improving teaching effectiveness. The Course Evaluation form was last revised in 2008. We spent Fall 2014 – Spring 2015 analyzing survey results and determining how to disseminate them. We report the results here and plan to disseminate them to faculty through Mason’s weekly E-Files Newsletter.

Executive Summary

About one-third of faculty (705 out of approximately 2100) responded to the online survey. We did not collect information on the distribution of full-time vs. part-time faculty from among respondents.

The survey consisted of eight items, a combination of multiple-choice and open-ended items. Analysis of faculty responses to these items revealed the following:

1) Fewer than half (40%) of all respondents were satisfied with the current course evaluation form, and almost one-third expressed dissatisfaction;

2) Almost two-thirds (64%) of respondents indicated some level of usefulness of evaluation results for improving course design and teaching;

3) Almost half (47%) of respondents indicated that free or open-ended responses were of most use to them;

4) Almost three-fourths (73%) of respondents said the course evaluation results were used for improving teaching;

5) More than half (55%) of all respondents indicated that course evaluations work in their favor;
6) Suggestions for making course evaluation forms more useful included putting the forms online, using fewer questions, using more open-ended questions, and adding new categories of questions;

7) Suggestions for eliciting more responses to open-ended items included using different types of questions, more open-ended questions, and more instructor-generated and course-specific questions;

8) Additional categories recommended for the course evaluation form were questions on use of Blackboard and technology by the instructor and using a different evaluation form for online courses and distance education. Respondents also suggested questions on student opinions on usefulness of course materials.

**Recommendations**

To address faculty concerns with the current Course Evaluation form and to increase the usefulness of the form for improvement of teaching effectiveness and course design and knowing that revising the Course Evaluation form is most likely a multi-year, iterative process, the Committee recommends the following actions:

1) Establish a university-wide competition for faculty and students to draft a new course evaluation form, pilot it, and validate it; or

2) Engage staff with survey development expertise in the Office of Institutional Assessment in developing and validating a course evaluation form; or

3) Purchase a proprietary course evaluation form that has already been validated.

4) Once developed and validated, the new Course Evaluation form should be examined for validity and reliability every three years by faculty with psychometric expertise, in conjunction with this Committee.

Developers of the new Course Evaluation Form would work with the Effective Teaching Committee and follow guidelines established by the Committee, including:
• Reducing number of Likert-scale questions and increase number of open-ended, course-specific questions, putting these on the front side of the form (rather than on the back);
• Adding categories such as effective use of technology or Blackboard;
• Making online evaluations available (on Blackboard) at discretion of faculty members and improving dissemination of information on availability of online evaluations;
• Using a specially-designed course evaluation form for online courses and distance education.

In addition, this Committee recommends a review of the research literature on the effectiveness of course evaluation forms and how they are used for faculty evaluation, promotion, and tenure. We have attached a starter list of references that can inform us as to the most recent research conducted on this topic.

This Committee is also providing a sample revised course evaluation form at the end of this report based on direct input from faculty on the survey and Committee members’ preliminary research on this topic. As we work to develop or purchase a validated evaluation instrument, we see the need to provide a modified course evaluation form based on faculty feedback. We provide specific examples of item revisions on the sample form. While these revisions are not yet empirically supported, they do reflect the concerns of faculty survey respondents. We would like to offer these revised items as an immediate response to a problem that calls for a long-term, research-based solution. We encourage the Faculty Senate to pursue more rigorous forms of survey development based on the recommendations listed above.

Effective Teaching Committee Members

Lorraine Valdez Pierce, Chair, College of Education & Human Development

Elizabeth DeMulder, College of Education & Human Development (Faculty Senator)

Danielle Rudes, Criminology, Law, & Society

Rodger Smith, Communication

Alexandria Zylstra, School of Business

April 14, 2015
Survey Results

1. To what extent do results from course evaluations help you improve your course design and/or teaching?

   a. A great deal
   b. Somewhat
   c. Not much
   d. Not at all

Number of respondents: 696

Almost half of all respondents (47%) indicated that course evaluations were somewhat helpful and an additional 17% indicated that they were very helpful, indicating that approximately two-thirds (64%) of respondents thought the course evaluation forms were useful to some extent in helping them improve course design or teaching.
2. Please rate the following categories in terms of their degree of usefulness for improving your course design and/or teaching:

[Likert Scale for each item below: (5) Very helpful, (4) somewhat helpful, (3) neither helpful or not helpful, (2) not helpful, (1) not at all helpful]

a. Course Organization and Planning  
b. Communication and Faculty Student Interaction  
c. Assignments, Exams and Grading  
d. Course Delivery  
e. Overall Rating of Teaching  
f. Overall Rating of Course  
g. Student Free Response

**Number of respondents: 701**

Responses indicated that all categories are somewhat helpful, but almost half of respondents (47%) indicated that student free responses were the most helpful to them.
3. How do you use ratings results obtained on Course Evaluation forms?

Select all that apply:

a. To improve your teaching  

b. For tenure, promotion, or contract renewal  

c. For program or department annual evaluation requirements  

d. For portfolio reflections  

e. To create your own benchmarks or performance goals  

f. For identifying which courses need more reflection and preparation  

g. To identify areas of strength and weakness  

h. Other [text box provided for written response]

Number of respondents: 682

Approximately three-fourths (73%) of respondents said they used course evaluation results for improving teaching, with almost two-thirds (62%) indicating that they used results to identify strengths and weaknesses in their teaching. Almost half of respondents indicated that course evaluation results were used for promotion, tenure and renewal (43%) or for annual evaluation requirements (44%).

| How do you use ratings results obtained on Course Evaluation Forms? Select all that apply. |
|---|---|---|---|---|---|---|---|
| To improve your teaching | For tenure, promotion or contract renewal | For program or department annual evaluation requirements | For portfolio reflections | To create your own benchmarks or performance goals | For identifying which courses need more revision or prep | To identify areas of strength and weakness |
| Series1 | 73.2% | 43.4% | 44.1% | 16.9% | 32.1% | 42.7% | 61.6% |
4. Do course evaluation results work in your favor for contract renewal, promotion and/or tenure decisions?

a. Yes  
b. No  
c. Not applicable to my situation

Number of respondents: 700

More than half of all respondents (55%) indicated that course evaluation results work in their favor. On the other hand, approximately one-third (32%) of all respondents indicated that course evaluations are not applicable to their situation. These may be adjunct instructors whose evaluation results are not used for making tenure-track decisions. However, without knowing the full-time or part-time faculty employment status of each respondent, we are unable to reach a definitive conclusion on this response.
5. How might course evaluation reports be structured differently to make them more useful to you (i.e., available online, fewer questions, available via phone app, new/revised categories)? [text box provided for written response]

Number of respondents: 705

Of the 596 respondents making suggestions for restructuring course evaluation forms, almost half (42%) recommended making these forms available online, more than a third (37%) recommended changing the type of question used, and 10% suggested adding more faculty- or course-specific questions.
6. What suggestions do you have for eliciting more responses to open-ended items on the Course Evaluation form? [text box provided for written response]

Number of respondents: 705

More than one-third of respondents (35%) did not provide any suggestions, but of the 460 responses that were provided, 43% indicated a need for different types of questions such as fewer Likert scale questions, more open-ended questions, instructor-generated and course-specific questions.
7. What categories or items, if any, would you like to see added to the Course Evaluation Form? (e.g., include other questions regarding type of course, Blackboard, technology, materials)? [text box provided for written response]

Number of respondents: 411

About one-fourth (26%) of respondents did not suggest additional categories, but of the three-quarters that did, 38% suggest questions on how well the instructor used technology (such as Blackboard) or about online courses, and about 25% suggested getting information from students such as their opinions on the value of the course and their perceived level of engagement in the course. 19% suggested adding more course-specific questions, such as the usefulness of materials.

Question # 7 - What categories or items, if any, would you like to see added to the Course Evaluation Form?
8. To what extent are you satisfied with the current Course Evaluation Form?

a. Very satisfied  
b. Satisfied  
c. Neither satisfied nor dissatisfied  
d. Dissatisfied  
e. Very dissatisfied  
f. Other Comments  [text box provided for written response]

**Number of respondents: 694**

Fewer than half (40%) of respondents indicated satisfaction with the current course evaluation form and almost one-third (28%) expressed dissatisfaction with the form. One-third (33%) of respondents had no opinion.
Sample Revised Items for Course Evaluation Form

Questions about the Instructor

1) Course requirements were clearly stated in the syllabus.

2) The instructor covered the important aspects of the course as outlined in the syllabus.

3) The instructor explained the material clearly.

4) The instructor encouraged students’ active participation through discussion and other activities.

5) The instructor used technology (such as Blackboard) effectively.

6) The readings helped me understand the material.

7) The course exams and/or required projects and assignments (projects, papers, presentations, etc.) reflected the content of the course.

8) Assignments, exams, projects, or other course requirements were returned in a reasonable amount of time.

9) Comments and suggestions on returned assignments (projects, papers, etc.) were helpful.

10) The instructor was accessible either in person or electronically.

11) The instructor responded to inquiries within a reasonable amount of time.

12) The instructor’s grading policy was clearly stated in the syllabus and followed.

13) The instructor showed respect for students.

14) What worked in this course and why? (could put this question first)

15) What suggestions would you make for improving this course next semester? (could put this question second)
Questions about the Student

16) The number of times you were absent from class: 8 or more 6-7 4-5 2-3 0-1

17) The grade you expect in this class: A B C D F

18) Your class level: Freshman Sophomore Junior Senior Graduate

19) Your overall GPA: 3.51-4 3.01-3.5 2.51-3 2.01-2.5 below 2.0

20) This course is: Required Elective Gen Ed Other

Suggested revisions to Course Evaluation Form by Effective Teaching Committee

April 14, 2015
Rationale for Suggested Revisions to Course Evaluation Form

1. **Re-ordered questions**: The committee suggests the survey might be most useful if the questions flowed in a chronological order. Thus, the re-ordered questions above reflect the chronology of a typical course: questions about the syllabus, instructor style, course content, and grading.

2. **Reducing the number of multiple-choice questions** to ten will leave more room for open-ended questions and questions specific departments wish to add to the form, as was requested by many survey respondents.

3. **Recommendation that questions 15 and 16 (the “overall” rating of the course and instructor) be removed** is based on research suggesting that they are ineffective and inaccurate measures of teaching effectiveness lacking in validity.

While student evaluations are useful for assessing observable information (*did the instructor follow the syllabus, were the readings helpful, was instructor feedback helpful, did the instructor follow the grading policy*), research suggests that due to different types of bias, student evaluations are not valid or useful for assessing overall teaching effectiveness. In a recent article evaluating student evaluations researchers concluded that:

- Student Evaluations of Teaching (SET) do not measure teaching effectiveness.
- Controlled, randomized experiments find that SET ratings are negatively associated with direct measures of effectiveness. SET seems to be influenced by the gender, ethnicity, and attractiveness of the instructor.
- Summary items such as “overall effectiveness” seem most influenced by factors not relevant to teaching effectiveness.

See annotated bibliography of Selected Readings on University Course Evaluations below, including Stark, P.B. & R. Freishtat (Sept. 2014).

This paper analyzed gender biases in over 22,000 undergraduate student evaluations of teachers (SETs) in a French university. Results suggest that male teachers tend to receive higher SET scores because of students' gender biases. Male students, in particular, expressed a strong bias in their favor: male students were approximately 30% more likely to give an excellent overall satisfaction score to male teachers compared to female teachers. The teaching dimensions for which students tend to perceive a comparative advantage for women (such as course preparation and organization) tend to be more time-consuming for the teacher, compared to the teaching dimensions that students value more in men (such as class leadership skills). Men are perceived as being more knowledgeable and obtain higher SET scores than women. Finally, if women increased students' continuous assessment grades (cumulative course grade before final exam) by 7.5% compared to the grades given by their male colleagues, they could obtain similar overall satisfaction scores as men. Yet, women do not act on this incentive (men and women give similar continuous assessment grades), suggesting that female teachers are unaware of students' gender biases. These biases have strong negative consequences for female academics, who may spend more time on teaching to try to obtain high SET scores, reducing time available for research, while SET scores do not seem to measure actual teaching effectiveness.


This paper contrasts measures of teacher effectiveness with students’ evaluations for the same teachers using administrative data from Bocconi University (Italy). Effectiveness measures were estimated by comparing the subsequent performance in follow-on coursework of students who were randomly assigned to teachers in each of their compulsory courses. Researchers
found that, even in a setting where the syllabi were fixed and all teachers in the same course presented exactly the same material, teachers still mattered substantially. The average difference in subsequent performance between students who were assigned to the best and worst teachers (on the effectiveness scale) was approximately 43% of a standard deviation in the distribution of exam grades, corresponding to about 5.6% of the average grade. Additionally, researchers found that the measure of teacher effectiveness was negatively correlated with the students’ evaluations: in other words, teachers who were associated with better subsequent performance received the worst evaluations from their students. Results were rationalized with a simple model where teachers could either engage in real teaching or in teaching-to-the-test, the former requiring higher students’ effort than the latter. Teaching-to-the-test guaranteed high grades in the current course but did not improve future outcomes. Hence, if students were myopic and evaluated better teachers from which they derived higher utility in a static framework, the model was capable of predicting the empirical finding that good teachers receive bad evaluations, especially when teaching-to-the-test is very effective (for example, with multiple choice tests). Consistent with the predictions of the model, researchers also found that classes in which high skill students were over-represented produce devaluations that were less at odds with estimated teacher effectiveness.


Multiple studies show that formative assessment can improve student learning in the classroom. The most in-depth study was conducted in 1998 by scholars Paul Black and Dylan Wiliam, who discovered that improvements in student learning through formative assessment “are amongst the largest ever reported for educational interventions” (Black & Wiliam 1998, as cited in Popham 2008, 2). However, less attention has been paid to an equally important phenomenon: how assessment through student course evaluations at midsemester influences students’ learning and end-of-semester course evaluation ratings. In the spring of 2012, University of Nevada, Las Vegas’ vice provost for
academic affairs appointed a Course Evaluation Task Group (CETG) to review UNLV’s disparate course evaluation approaches and recommend alternatives. During the process of considering alternative solutions, CETG members began discussing how their use of formative assessments at midsemester could improve student learning and end-of-term summative course evaluation results. Results obtained by UNLV instructors clearly showed that both student learning and classroom performance improve when instructors listen to their students’ suggestions and consider implementing them to improve teaching effectiveness. The evidence also shows how instructors receive higher end-of-course ratings.


A response to Schonrock-Adema et al.¹ who argue for use of a prediction-based method of course evaluation (i.e., students are asked to estimate the opinions of their classmates) rather than giving their own opinion. Dolmans et al. question the premise that students’ ratings of teaching are influenced by bias (e.g., grading leniency) and also express concern about time and effort that the prediction-based evaluation takes. The researchers conclude that when evaluating courses or teachers, both qualitative and quantitative data collected from various sources and stakeholders should be used. ¹Schonrock-Adema J, Lubarsky S, Chalk C, Steinert Y, Cohen-Schotanus J. ‘What would my classmates say?’ An international study of the prediction-based method of course evaluation. Med Educ 2013;47:453–62.


In Study 1, students received instructor specific grades and then took a common high stakes post-test that was centrally graded, thereby serving as a basis for measuring student learning. While small in magnitude, there was a
robust positive association between student learning and course evaluations. Strongest predictors of learning were questions about clarity of the instructor, use of supplementary materials and overall course experience. However, students’ grades in the course were more strongly correlated with their course ratings than the learning was. In Study 2, patterns of nonparticipation in online course evaluation of 20,000 students were examined. Overall response rate for course evaluations was about 50 percent. Patterns of nonresponse were not random: Lower grades, courses not in the major, social majors (as opposed to “realistic majors” like chemistry, biology), and survey fatigue were factors related to nonresponse. (1.) Marks, M., Fairris, D. & Beleche, T. (2010) Do Course Evaluations Reflect Student Learning? Evidence from a Pretest/Posttest Setting. Working paper: http://faculty.ucr.edu/~mmarks/Papers/marks2010course.pdf (2.) Adams, M. Umbach, P. (2010) "Who Doesn’t Respond and Why? An Analysis of Nonresponse to Online Student Evaluations of Teaching" presented at the annual meeting of the Association for the Study of Higher Education.


Reports on emerging trends (online evaluation tools, connecting evaluation data to accountability measures and competency-based learning outcomes, increasing use of evaluations for formative purposes, and contextualization of evaluation data for summative evaluation of teaching) as well as existing gaps and recommendations for further research (defining teaching vocabulary and expectations, understanding evaluation users, educating evaluation users, and evaluating graduate student teaching assistants and instructors).

Jones, S. J. (2012). Reading between the lines of online course evaluations: identifiable actions that improve student perceptions of teaching effectiveness and course value. Journal of Asynchronous Learning Networks, 16 (1).

This paper examines whether online course evaluations can provide ways for faculty to improve teaching and effectiveness. The study was conducted at a large research university in US with 30,000 students. Researchers collected 2,826 online evaluations (23 questions; Likert Scale) between 2007-9. Results suggest that the two most important factors are: 1) Stimulation of learning (high quality,
rigorous, engaging, organized) matters most in perception of teaching effectiveness and 2) usefulness/relevancy of assignments important to students' perception of course. Also, clearly presenting information was deemed to be important and scored highly.


Researchers took two online course instructors, one male and one female, and gave them two classes to teach. Each professor presented as his or her own gender to one class and the opposite to the other. Students gave professors perceived to be male much higher evaluations across the board than they did professors perceived to be female, regardless of what gender the professors actually were. Students penalized the perceived female professor in all 12 categories, including in qualities that women are usually assumed to excel at, such as being caring and respectful. This comports with other studies that show that while female professors are judged somewhat less harshly if they conform more to female stereotypes, men still get bonus points for showing up male.


This study asks whether class and teacher ratings are related to attendance and individual student differences. Additionally, it considers whether online evaluations have a lower response rate than paper evaluations. The researcher collected data from students in all statistics class evaluations in 2009 (fall) in a large university. The unit of analysis was class selection, not students. The study also used some random assignment in some course sections where half of the class were given online evaluations and the rest received paper forms. All distance courses only used online and rest (non-randomly assigned) were given choice of paper or online evaluations. Data included 25 online and 11 paper evaluations from graduate students and 36 online and 9 paper evaluations from undergraduates. The overall and statistically significant finding from this study reports that online evals do not negatively impact course/instructor ratings and
response rates. Also instructor and course ratings do not depend on student attendance and undergraduate students do not have lower variability than grad students.


This paper explores the factors that predict students' participation in online evals. Researchers collected 2,478 evaluations from 895 (undergraduate and graduate students) in 24 courses during 2010-11 at one university with a 26.8% response rate. The survey sampling design lent itself to multilevel analysis (questionnaire level, student level, course/teacher level). Findings suggest that at the questionnaire level, student course grades are statistically significant predictors of students' completing evaluations. At the student level, the number of evaluations they receive and their education level (maturity) are both statistically significant predictors for course evaluation. Nothing was statistically significant at the course/teacher level.


Student ratings of teaching have been used, studied, and debated for almost a century. This article examines student ratings of teaching from a statistical perspective. The common practice of relying on averages of student teaching evaluation scores as the primary measure of teaching effectiveness for promotion and tenure decisions should be abandoned for substantive and statistical reasons: There is strong evidence that student responses to questions of “effectiveness” do not measure teaching effectiveness. Response rates and response variability matter. Comparing averages of categorical responses, even if the categories are represented by numbers, makes little sense. Student ratings of teaching are valuable when they ask the right questions, report response rates
and score distributions, and are balanced by a variety of other sources and methods to evaluate teaching.


Available online at [http://www.biomedcentral.com/1472-6920/15/30](http://www.biomedcentral.com/1472-6920/15/30)

For the purpose of this systematic review, online databases (PubMed, PsycInfo and Web of Science) were searched up to August 1st, 2013. Original research articles on the use of student ratings in course evaluations in undergraduate medical education were eligible for inclusion. Twenty-five studies met the inclusion criteria. Qualitative research (2 studies) indicated that overall course ratings are mainly influenced by student satisfaction with teaching and exam difficulty rather than objective determinants of high quality teaching. Quantitative research (23 studies) yielded various influencing factors related to four categories: student characteristics, exposure to teaching, satisfaction with examinations and the evaluation process itself. Female gender, greater initial interest in course content, higher exam scores and higher satisfaction with exams were associated with more positive overall course ratings.


In two subsequent experiments, the influence of mood on academic course evaluation is examined. By means of facial feedback, either a positive or a negative mood was induced while students were completing a course evaluation questionnaire during lectures. Results from both studies reveal that a positive mood leads to better ratings of different dimensions of lecture quality. With regard to course evaluation, it can be expected that negative mood might cause students to focus more on criticizing details of a course while positive mood might lead to a global and mostly positive overall impression.