

Faculty and Student Perspectives on Student Teaching Evaluations

This study aimed at exploring faculty and student perspectives on student evaluation, as well as identifying their perceptions of the usefulness and appropriateness of the ratings for evaluating teaching effectiveness. More specifically, the study liked to identify the consequences, both intended and unintended, of using the evaluations, in addition to better understanding the process students use in responding to evaluations and what use faculty make of the evaluations. Through an analysis of consequences, institutions may learn how to maximize the positive while avoiding the negative, thus improving the measures, the measurement process, and the validity of inferences made (Ory & Ryan, 2001).

Method

Tool

The Instructor Course Evaluation (ICE) forms have been in use at the University for the past three years and have undergone extensive reliability and validity. However, there is a need to explore student and faculty input and perceptions of the process, especially as several exchanges revealed faculty dissatisfaction with certain aspects of the administration. In addition, it was important to identify the seriousness with which students approached the process in order to support the validity of the assessments. The main components that need to be surveyed were identified through a review of the literature and the questions raised by students and faculty. Based on them two survey forms were developed, one for faculty and the other for students. In order to assure respondent motivation, care was taken to make survey forms short, clearly worded and at the same time answering questions raised. Respondents were to answer items using a 5-point scale ranging from strongly agree to strongly disagree, in addition to providing comments on the system and the process.

The student survey included 12 items covering the following components: a) student perception and use of the process; b) purposes they think ratings are used for; c) consequences (positive and negative) of the use; d) motivation and seriousness of students while filling forms; and e) meaning of the 5-point scale to them. In addition, some demographic information like grade level and faculty were requested. (Table 2)

The faculty survey included 12 items and covered following components: a) purposes for which ratings are used; b) positive and negative consequences of use; c) faculty perceptions of the process; and d) attitudes towards being evaluated by students, in addition to some demographic items like academic rank and faculty (Table 2).

Sample

All students and faculty in five faculties at the University participating in the Instructor Course Evaluation (ICE) system made up the population of the study. At the beginning of the fall term 2005-6, e-mails were sent to all students and faculty soliciting their cooperation and requesting their input on the ICE system through answering the questions raised in the surveys. The surveys were posted on OIRA (Office of Institutional Research & Assessment) website and administered electronically. They were kept for three weeks on website during which faculty and students were reminded of the survey and the need to fill it out. At the end of the period, 605 students and 145 faculty members have filled out the surveys. Breakdown of the sample by faculty, rank (faculty), and class

level (students) is provided in Table 1. An examination of the Table shows that the sample is representative of various groupings and of the University population.

Data Analysis

Data were analyzed using SPSS. Frequencies and descriptive information were reported per item. Faculty and student perceptions were compared qualitatively and statistically through investigating mean differences using t-tests of uncorrelated data. Furthermore differences in responses by faculty, rank, and class level were investigated using analysis of variance.

Table 1. Breakdown of the Sample by faculty, Rank, and Class Level

Faculty	Faculty			Students		
	No	%	AUB%	No.	%	AUB%
Arts & Sciences	74	51	53	236	39	40
Agriculture & Food Sciences	12	8	8	52	9	9
Engineering & Architecture	21	15	19	130	22	24
Health Sciences	13	9	6	43	7	5
School of Business	17	12	11	122	20	20
School of Nursing	3	2	3	14	2	2
Missing	4	3		8	1	
Total	144	100		605	100%	
Rank	No.	%	AUB%			
Full Professor	27	19	23			
Associate Professor	20	14	16			
Assistant Professor	49	34	35			
Lecturer	14	10	7			
Instructor	31	22	18			
Missing	3	2				
Total	144	100				
Class level				No.	%	
Graduate				113	19	
Senior				169	28	
Junior				152	25	
Sophomore				153	25	
Freshman				9	2	
Missing				9	2	
Total				605	100%	

Results

Student Perceptions

Results of student responses to survey are reported in Table 2. Item means, medians, percent agree (%A) and disagree (%DA) are provided. With respect to uses made of ICEs, 70% of students perceive them to be a means for indicating suggestions for improvement and around half of them believe that faculty value input from them and make improvements as a result of weaknesses identified. Moreover, around two thirds of

students **perceive** evaluations as an effective means of evaluating faculty and do not agree that the ratings are a meaningless activity. As to **consequences**, around half of the students do not believe that faculty alter their teaching to get high evaluations, but at the same time believe that the content of the rating form affect what is addressed in class. A very high percentage of students perceive the rating **process** as allowing honest evaluations, and state that they are giving adequate thought and effort to it, in addition to being fair and accurate in their evaluations. However interestingly, when asked about their peer's attitude, opinion is equally divided between agree, disagree, and uncertain. As to the **5-point rating scale**, it seems it is well-understood by students with around half of them opting for a '3' when they are undecided or uninterested. However, another 40% disagree with their colleagues with regard to the use of '3'.

Lowest item mean rating went to 'Ratings are a meaningless activity that is performed by students and faculty only because it is required' and to 'and to 'Instructors alter their teaching in order to receive high ratings (give high grades, lower content difficulty, provide less content)'. While, highest ratings went to 'You are fair and accurate in your ratings of teachers and give adequate thought and effort to the rating process' and to '. The current procedure for rating faculty sufficiently protects your identity and allows you to be as honest as you would like to be'.

Faculty Perceptions

Faculty responses to survey questions are also reported in Table 2 in terms of descriptive statistics and frequencies. Majority of faculty value input from ratings and make **use** of them to improve their courses based on weaknesses identified by students. However, only a few of them agree that they should be used for making personnel decisions regarding salary and promotion. As to **consequences** of the evaluations, more than half of faculty respondents believe that faculty change their teaching to receive high ratings, while around 40% assert that what is addressed in class may be determined by content of ratings. Nearly half of the faculty do not believe that the ratings result in negative consequences like reducing faculty morale and job satisfaction .With respect to their **perceptions** of the ratings, around half disagree that they are a 'meaningless activity' while a third agrees with that statement. Majority of faculty believe that demanding a lot from students will result in lower evaluations, while there was a split of opinion with respect to 'Good instructors get high course evaluations'. As to faculty view of **process students** use to fill out the ratings, majority believe that students do not take evaluations seriously and around half perceive that students do not have enough knowledge to judge quality of instruction. However, most of faculty disagree with 'faculty members should not be evaluated by students'.

Highest faculty mean rating went to items 1 (Instructors value input and make improvements ..) and 17 (Instructor s who demand a lot from their students get low evaluations), while lowest evaluations (agreement) went to items 14 and 20.

Table 2. Descriptive Statistics and Frequencies of Student and Faculty Responses

Survey Items	Students (n=605)				Faculty (n=144)			
	X	Md	%A	%DA	X	Md	%A	%DA
Student								
*1. Instructors value input and make improvements in their teaching & courses as a result of weaknesses identified by students.	3.2	3.0	48	27	3.5	4.0	66	20
2. Students perceive and use ratings as a means for indicating suggestions for improvement	3.8	4.0	71	11				
*3. Instructors alter their teaching in order to receive high ratings (give high grades, lower content difficulty, provide less content).	2.7	2.0	23	51	3.3	4.0	54	32
*4. Content of student rating form may determine what is addressed in the classroom.	3.3	4.0	53	23	3.1	3.0	44	32
*5. Ratings are a meaningless activity that is performed by students and faculty only because it is required.	2.6	2.0	26	58	2.9	3.0	35	49
6. Rating forms provide an effective means of evaluating instructors.	3.5	4.0	64	21				
7. The current procedure for rating faculty sufficiently protects your identity and allows you to be as honest as you would like to be.	4.1	4.0	84	7				
8. You are fair and accurate in your ratings of teachers and give adequate thought and effort to the rating process.	4.3	4.0	93	3				
9. Other students are fair and accurate in their ratings of teachers and they give adequate thought and effort to the rating process.	2.9	3.0	27	31				
10. The 5-point rating scale is easy to use and understandable.	4.0	4.0	86	7				
11. A perfect five on rating scale is unobtainable.	2.8	2.0	31	51				
12. You usually use 3 in your ratings because of lack of interest or when you are unable to answer.	3.2	4.0	53	38				
Faculty								
13. Instructors are rewarded for having excellent rating results.					2.8	3.0	31	47
14. Most students take the evaluation process seriously					2.4	2.0	39	60

15. Students do not have enough knowledge to judge the quality of instruction.					3.0	3.0	41	48
16. Good instructors get high course evaluations.					3.0	3.0	40	33
17. Instructor s who demand a lot from their students get low evaluations.					3.6	4.0	60	23
18. Rating forms provide an effective means of making personnel decisions regarding salary and promotion.					2.4	2.0	19	58
19. Student ratings pose a threat to academic freedom and reduce faculty morale and job satisfaction.					2.8	3.0	32	46
20. Faculty members should not be evaluated by students					2.3	2.0	16	74

* Items on both surveys

Differences between Student and Faculty Perceptions

Responses on four common items on surveys (#s 1, 3, 4, 5) revealed significant mean differences between faculty and students ($t=2.89, p<.01$; $t=4.98, p<.01$; $t=2.19, p<.01$; $t=2.76, p<.01$, respectively). Two thirds of faculty believe that instructors value input versus around half of students. Greatest difference was noted on consequences of ratings in terms of faculty altering their teaching to receive high ratings, as most of the students did not believe that this was happening while majority of faculty perceived it to take place. Higher percentages of students perceive meaningfulness of ratings and believe that their content may determine what is addressed in class.

Differences in Ratings by Faculty, Rank, and Class

Investigating differences between student responses by class level, revealed significant differences on five items, 1, 5, 8, 9, and 11 ($F= 3.145, p< .014$; $F=3.811, p< .005$; $F=2.405, p< .049$; $F=4.114, p< .003$; and $F= 3.875, p< .004$, respectively) of survey. Scheffe post-hoc tests to identify source of differences revealed that it was always with sophomore students whose means differed significantly from others. As to differences in student responses by faculty, they were practically non-existent except on two items, 4 and 12 ($F=4.008, p< 0,003$; $F=3.835, p< .004$, respectively), where faculty of Agriculture and Food Sciences students had significantly higher means than others.

As to differences in faculty responses, no significant differences were found by rank and only one item, 3 ($F=2.590, p< .04$) showed differences by faculty. Post-hoc tests revealed that the difference was between Faculty of Health Sciences and School of Business faculty responses, where the later believed much less that ‘instructors alter their teaching...’.

Open-ended Comments

Both faculty and students provided extensive comments on the ICE. Qualitative analysis was conducted and the following present main points raised.

Faculty comments. They mainly focused on the following dimensions:

- a. Need to enhance communication about the system. In this respect students need to better understand process and respect it and faculty need to receive positive incentives. Conducting open forums or sessions to discuss importance of evaluations and to take stakeholder perspective is one method to enhance understanding and acceptance. New faculty need to be given special consideration in this regard.
- b. Need to revisit questionnaire as it was commented that i) wording of some items is not clear, and ii) items are not tailored or contextualized to particular needs of courses and accordingly information provided is not very useful.
- c. System needs to be expanded to include other means of evaluating faculty member performance. The current student evaluation system should be one of many available like peer review and others.
- d. Need to revise current ICE administration/collection procedures now in use as it is leading to major infractions and damage to the system.
- e. Faculty value evaluations and make use of them to make adjustments, especially the comments. Trend analysis recently reported over a number of semesters provided good indices of teaching effectiveness. In fact, most of those who provided input perceived the purpose of evaluations mainly for improvement and they were not to be used for promotion and personnel decisions unless they are based on trends and supplemented with other modes of evaluation.
- f. Faculty cited some biases associated with the system like ‘ratings affected by mid-term exam grades and difficulty of course’, ‘had to water down content of courses to make students happy and have fun’, ‘students want easy way out, etc...
- g. Although faculty agreed with the need to be evaluated by students, yet they questioned the seriousness with which students responded. For example, ‘students take evaluations seriously first time only, it becomes meaningless and mechanical later on’, ‘many of our students are not mature or responsible enough to give accurate evaluations’, ‘students evaluate their own achievement and not instructor’, ‘they are influenced by amount of work and grade’.

Student comments. Students provided extensive comments focused mainly on following issues:

- a. Recommendations were provided to revise form, make it shorter, and different from one major to another. In addition, administration should be done earlier in semester in order to give faculty member chance to adjust and students to see effect of evaluation. Many comments were provided on present system of administration in class with Graduate Assistants (GAs.) and this need to be addressed. Electronic submission was recommended as it better protects identity of students especially while writing comments.
- b. Students perceive that their evaluations are not made use of and accordingly do not take them seriously. Comments like ‘we need to feel that what we say matters’, ‘we students feel that evaluations are useless as not a single time

there is a consideration of our preferences’, I point out positive and negative points of instructor hoping that somebody listens to what I have to say however thinking that nobody really cares’, ‘I don’t think it matters what I write about the professor’, ‘most students fill form with inappropriate choices because they don’t believe that the surveys will do any good’, and ‘we never really know how ratings are used and if instructors themselves take them into consideration’ were cited.

- c. Students perceive evaluations to be important but that some students fill them inaccurately. Several reasons were cited for that like ‘lack of time’, ‘pressure from GA’s’, ‘students rather finish in haste and get out of class sooner’, ‘some professors make light of evaluations which leads us students to think that they do in fact make nothing’, and ‘personal likes and dislikes affect ratings’.
- d. Students called for better communication regarding the evaluations like ‘students must be informed about the importance of the surveys’ and ‘students should be told how and why the results of the ICE will be used’.
- e. Students strongly called for publicizing results of the evaluations. They felt that ‘it is better to register with a certain instructor because of numerical data rather than campus rumors’ and ‘students need to access their instructors’ evaluations, it is only when this process becomes transparent and understandable to students that the university’s gap of disenfranchised apathetic students may be bridged and the integrity of instructor quality may be credible’.

Conclusion and Recommendations

This study has attempted to study the purposes and actual uses of the instructor course evaluation system at the University, in addition to the perspectives of the different parties involved with respect to the usefulness and appropriateness of the ratings for evaluating teaching effectiveness. Results revealed that both faculty and students positively valued student evaluations especially for making improvements, believed that they were made use of for that purpose, and did not think that the process was meaningless. Both of them called for revision of the form currently in use, improving administration procedures, and enhancing communication between students, faculty, and administration on the system, its purposes, uses, etc. Faculty were concerned about ‘biases’ associated with student evaluations, with seriousness with which students filled the forms, and with their use for personnel decisions especially as a sole indicator of teaching effectiveness, and accordingly called for expanding the system to include other measures. Students were more concerned that faculty and administration were not taking their evaluations seriously and wanted evidence of making use of the evaluations, in addition to having results publicized.

This study has attempted to investigate the validity of the instructor course evaluation system at the University. The findings provided partial support as some aspects of validity were validated while others needed to be improved on. Accordingly, and in order to enhance the validity of the system it is recommended to take the following measures:

- a. Revision of the form and administration procedures currently in use in line with student and faculty feedback (wording of some items, frequency and time of administration, contextualization, training of GAs, etc.).
- b. Start a dialogue on the subject between students, faculty and the administration. There is a state of general tension surrounding course evaluations, which appears to influence instructor attitudes. Faculty are apprehensive of process used by students and at the same time feel that they are not rewarded for good teaching. So providing incentives for improving instruction and a support system to accompany it will help reduce this tension. Similarly, there is a need to work on student attitudes, to motivate them to participate in the process, and to convince them that their opinions do matter. Students cannot be expected to understand the importance of evaluations if they have no knowledge regarding how the evaluations will be used, or if teaching evaluations are used for purposes that students do not value, or if they see no visible results from their participatory efforts. A series of open discussion forums and dialogues will also help to enhance the process and reduce the tension.
- c. Expand the system to include other measures of teaching effectiveness especially for personnel decisions. McKeachie (1997) recommends adopting instructional goals attainment approach for summative evaluations used for personnel decisions. Students' ratings of their attainment of educational objectives not only provide better data for personnel committees but also stimulate both students and teachers to think about their objectives- something that is educational in itself. Similarly, to reduce 'bias' associated with factors unrelated to teaching effectiveness when making comparative decisions, it is important not to make crude comparisons of raw student ratings without any consideration of the instructional contexts in which those ratings are generated. Judgment or decisions based on raw student ratings will be likely at fault unless such potential biases are taken well into consideration in the use and interpretation of the information (Kwan, 1999). Trend analysis should be adopted and not single evaluations.
- d. Enhance transparency by having clear purposes and uses of the system, in addition to making results public.