Student evaluation of course and teaching: a case study

Jacek Uziak

University of Botswana
Gaborone, Botswana

ABSTRACT: The application of the student evaluation of courses and teaching (SECAT) to courses in the Department of Mechanical Engineering at the University of Botswana is discussed in this article. Response rate for SECAT for the two semesters reported in the article was low, as expected for a survey administered on-line. From the quantitative analysis it was proved that there was no correlation between the SECAT results and failure rate or average mark for the course. In terms of the qualitative analysis, students reported different preferences on course attributes for different types of course; but generally, students liked the most courses which were related to industry and engineering practice. Students also indicated the most important attributes of the lecturer's effectiveness, as well as attributes of a *good* and *bad* lecturer.

INTRODUCTION

Student evaluations of teaching (SETs), also called course evaluations, have formally existed since the 1920s [1][2]. The SET appears in different formats and may be referred to differently, such as student assessment of course and teaching (SECAT) or student evaluation of instruction (SEI), depending on the institution. Irrespective of the name or abbreviation, such assessments seem to be an almost universally accepted method of gathering information about the quality of education [3]. Evaluations may have multiple uses at various institutions, and there are some important decisions that typically follow from the results of the exercise:

- Administrators use teaching evaluations for annual review, promotion, tenure, and reappointment decisions [4].
- Department heads may consider results from evaluation to decide whether to maintain the course in the curriculum or to change it.

However, the evaluation also can be a great motivation for staff, who can use the results to improve their instructions and apply for grants or awards [5].

Universities generally treat students' ratings primarily as a measurement of teaching effectiveness or teaching quality. In particular, quantitative evaluations of instructors' overall teaching effectiveness frequently are emphasised in personnel decisions [6]. Although it may seem natural to think that students assess staff teaching effectiveness, it is at most not an obvious conclusion. Students in their assessment have biases that might influence the results [7].

Bias occurs when …*a teacher or course characteristic affects teacher evaluations, either positively or negatively, but is unrelated to criteria of good teaching, such as increased student learning* [6]. Worse, in some circumstances the association between evaluation and an objective measure of teaching effectiveness is negative [7]. One of the fundamental issues related to student assessment of teaching concerns the response rates. Although in data-driven decision making it is essential to gather the assessment representative to the whole population that rarely happens. A steady decline in survey participation has been observed due mainly to switching from paper to Web-based modes of administrating surveys [8][9]. Increase in non-response creates potential errors in any survey; hence, the quality of data and their results tend to weaken when response rates decline [9][10]. It is important to understand student non-response to on-line teaching evaluation. It is also important to understand non-response bias:

- Non-response increases the potential for error and is a threat to external validity.
- In most cases, survey non-response is not random [8][11][12]; that is, there is a reason for its occurrence.
- Low response rates are likely to increase bias in results, thus affecting the external validity of the results [10-15].
- As non-response increases, the likelihood that non-respondents' opinions differ from respondents' also increases [10][15-17].

MATERIALS AND METHOD

The student evaluation of courses and teaching (SECAT) has been a tool to assess teaching and courses at the University of Botswana (UB) for almost 20 years. It has not changed in the past 10 years and in recent years has been applied only on-line. The SECAT form contains sections on teaching evaluation (13 questions) and course evaluation (11 questions). An open-ended question section on general appraisal requests students to indicate what they liked most about the lecturer and the course, and to explain low ratings in teaching and course evaluation sections. It also has one question to assess *lecturer's overall effectiveness*. A special interest in that question is related to the fact that for years the assessment of staff teaching was based purely on that question. However, that has changed and in the past few years the teaching assessment is based on an overall average of all questions in the quantitative parts of the questionnaire, i.e. teaching and course evaluation. The numerical SECAT result is one of the main tools to assess staff teaching. The exercise is done every year for all staff.

The analysis presented in the article is based on data collected for two semesters; 2016/2017 - semester 2 and 2017/2018 - semester 1, in one of the departments in the Faculty of Engineering and Technology at UB. The data included cover all lecturers and courses in mechanical and industrial engineering, two programmes offered by the department. The assessment covered also general education courses (GEC), which constitute elements of general education at UB.

The response rate for SECAT was low, in semester 1 - 60%, whereas in semester 2 it was even lower - 36%. There is a clear explanation of that low response rate; it is only expected as SECAT was administered on-line with no control on whether the students participated.

RESULTS

Quantitative Analysis

Both teaching and course averages were relatively high, and consistent in both semesters (Figure 1). There was a slightly better assessment of teaching than courses, which was more visible in semester 2. Figure 1 also shows pooled standard deviations; for teaching 0.78 and 0.82 for semesters 1 and 2, respectively, and 0.91 and 0.95 for courses in semesters 1 and 2, respectively.

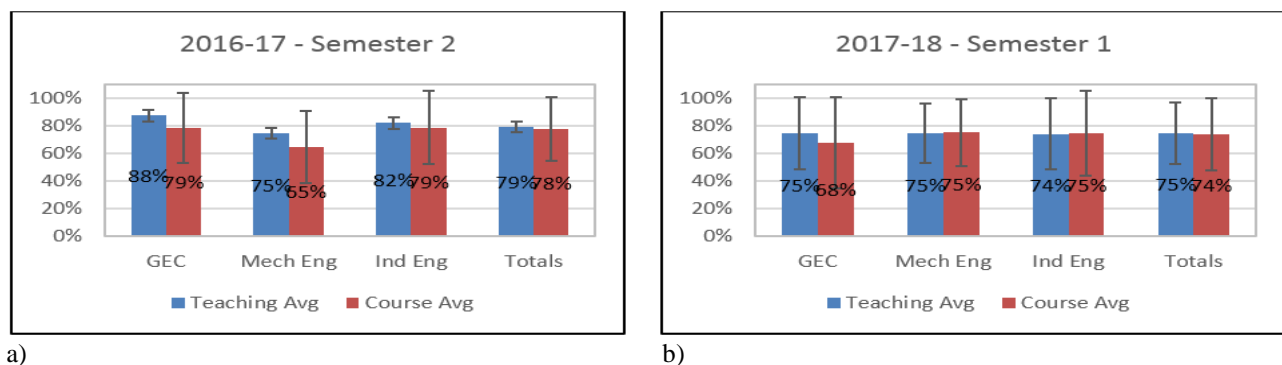a)   b)

Figure 1: SECAT averages; a) Semester 2 and b) Semester 1.

Correlation analysis indicated no relationship between:

- SECAT teaching average and failure rate (coefficient of correlation for semester 2 - 0.24622, and for semester 1 - 0.24504);
- SECAT teaching average and average student mark (coefficient of correlation for semester 2 - 0.288095, and for semester 1- 0.30933);
- SECAT course average and failure rate (coefficient of correlation for semester 2 - 0.06963, and for semester 1 - 0.2848);
- SECAT course average and average student mark (coefficient of correlation for semester 2 - 0.058491, and for semester 1 - 0.300012).

Qualitative Analysis

Quantitative analysis for SECAT was done by investigating students' answers to open-ended questions. Such analysis can give insight into contextual understanding and understanding interactions; all elements very important in the

educational environment. The open-ended questions, and the number of answers in different type of courses, i.e. general education courses (GEC), Industrial Engineering (IE) and Mechanical Engineering (ME), are presented in Table 1.

Table 1: Questions and number of answers.

Questions	Number of answers			
	GEC	IE	ME	Total
What did you like most about this course?	36	181	551	768
What did you like most about your lecturer?	36	182	558	776
Identify the questions rated low and explain why they were rated as such.	11	36	88	135
Explain why you gave the rating for your lecturer's overall effectiveness.	34	150	444	628

Course Appreciation

Overall, students liked most when a course was related to industry and actual practice (see Figure 2). However, there were differences between different types of course (Figure 3). The importance of the relationship between the course and industry and practice was also confirmed by specific comments such as:

- *He prepared us not just for the tests but also for work in industry.*
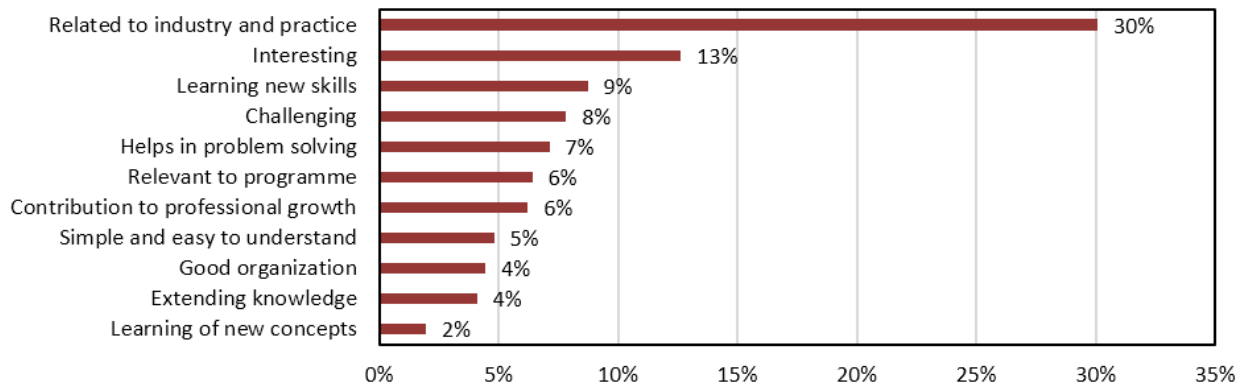- *The assignments given were very practical and industry related.*

Figure 2: Overall course appreciation by students: what did you like most about this course?

For mechanical and industrial engineering courses, their relation to industry and practice was indeed the most important aspect. However, for the general education courses (GEC) it was the contribution of the course to professional growth that was most important, which had little importance for the mechanical engineering courses. Equally important for GEC was good organisation, again not essential for ME and insignificant for IE. There was also a dramatic difference in *learning new concepts* and *learning new skills.* For both ME and IE courses, skills learning was relatively important, but learning new concepts was not significant for GEC (Figure 3).
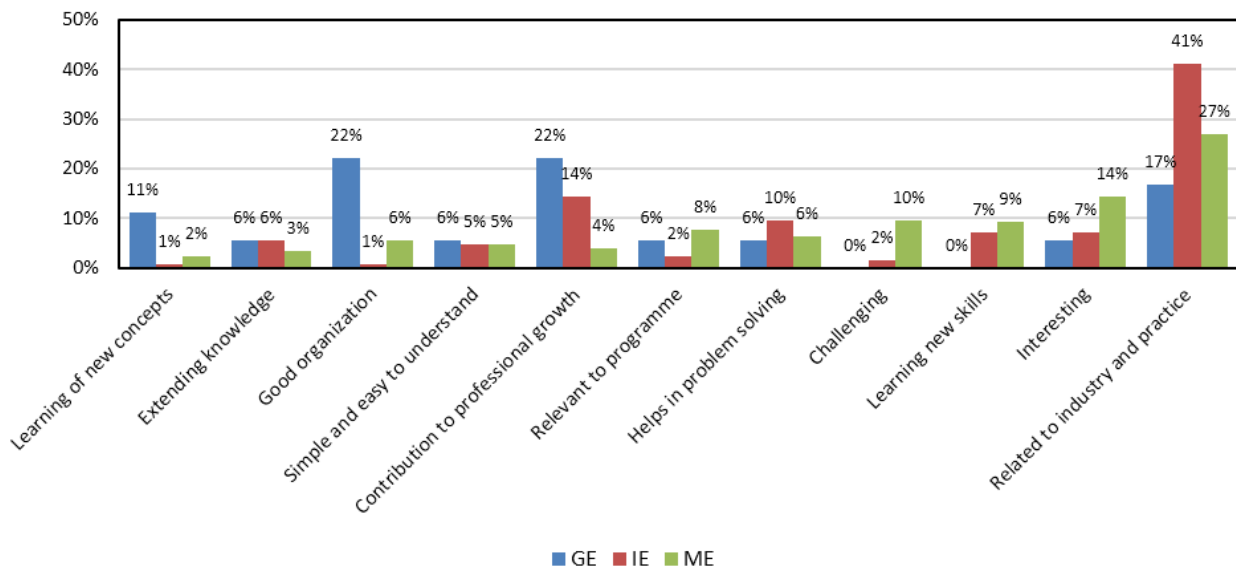

Figure 3: Course appreciation by course type: what did you like most about this course?

Other significant factors identified from the open-ended questions included exposure to general knowledge (GEC), learning life skills (IE courses), personal growth (IE courses) and fun (ME courses). However, the number of such answers was not significant.

Lecturer Appreciation

The second open-ended question provided the students with an opportunity to specify their *model* lecturer (Figure 4). Explaining concepts was the most important factor for students, followed by enthusiasm, class management to provide a conducive atmosphere, and willingness and availability to assist. In a society plagued by tardiness, the high value given to punctuality is worth noting.

The other striking information is the very low student appreciation of lecturer knowledge of subject, and linking theory and application. The latter is in direct contrast to course considerations, where relating to industry and practice was considered the most important aspect. However, the slight endorsement of subject knowledge is quite significant unless the students thought that it would be difficult for them to assess it. Sadly, there is also an indication that some students did not appreciate their lecturers at all, indicating that they liked *nothing* about them.
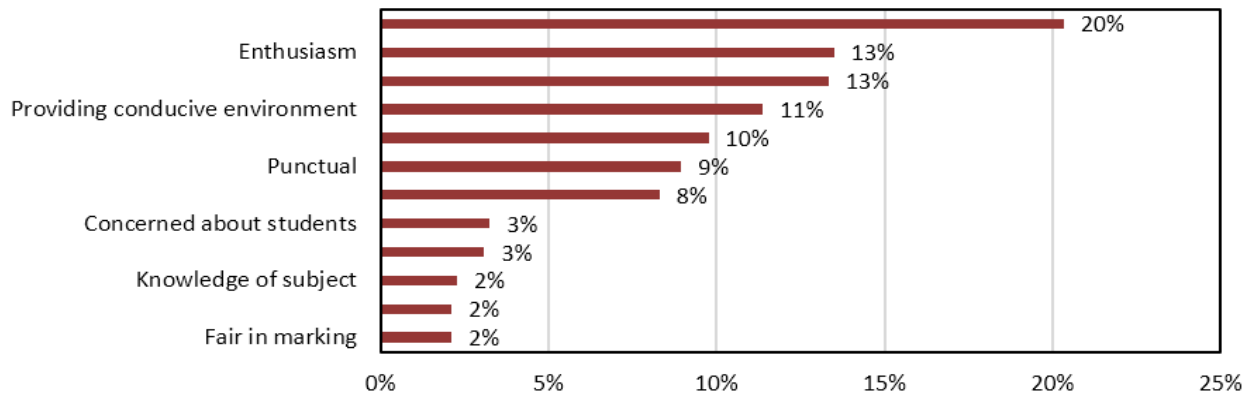

Figure 4: Lecturer appreciation: what did you like most about your lecturer?

The other listed factors, although not gaining much popularity, were taking the job seriously, being charismatic, presentable and being loud to be heard.

Explanation of Low Ratings

The analysis of low ratings gives an insight into students' perception of the symptoms of a bad lecturer (Figure 5). The claim of not fair marking or marking not done on time was indicated as the most common indicator. It was followed by clear elements of pedagogy, such as bad delivery, no use of small group teaching, not providing hand-outs and the lecturer being disorganised. Course characteristics such as being hard and demanding or not interesting were also recognised, although at a lower rate.
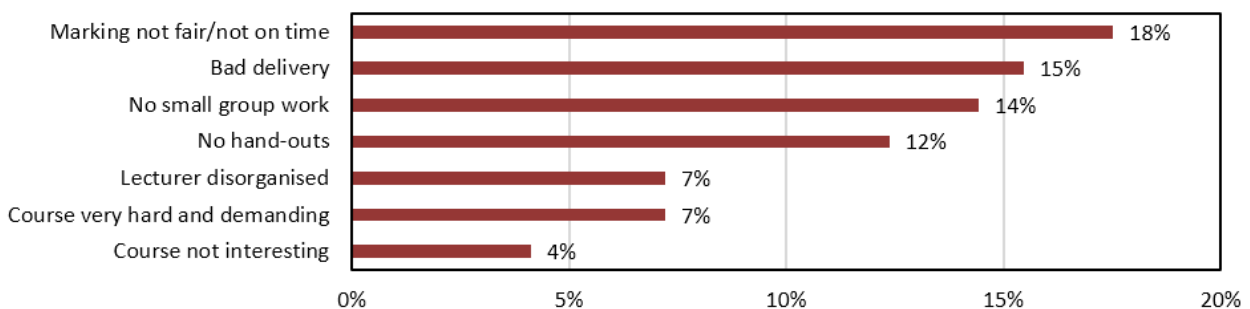

Figure 5: Explanation of low ratings for both course and lecturer.

Students provided both positive and negative comments supporting their general explanations. For markings they claimed that some lecturers are:

- *Not fair in marking the assessments;*
- *Does not mark fairly;*
- *Not consistent in marking.*

Teaching also has been commented on in negative ways indicating the importance of clear delivery, which should help to understand the concepts:

- *Has never taught us/did not teach or explained/he only guides us;*
- *His lecturing was boring. He did not show any enthusiasm about the course;*
- *Too fast, did not explain concepts clearly.*

Additional explanation, though only occasional, included complaints that the course was preparing only for tests and examinations, lecturer being rude, very strict or completely unavailable outside of classroom.

Attributes of Lecturer's Effectiveness

The most commonly listed attributes of lecturer's effectiveness are pedagogical elements, such as coverage of course material, course management and the ability to explain concepts (Figure 6). Additional factors included creating a conducive atmosphere in the class and engaging students in the teaching/learning process. Further, enthusiasm for the subject and teaching in general, and punctuality again, are recognised.
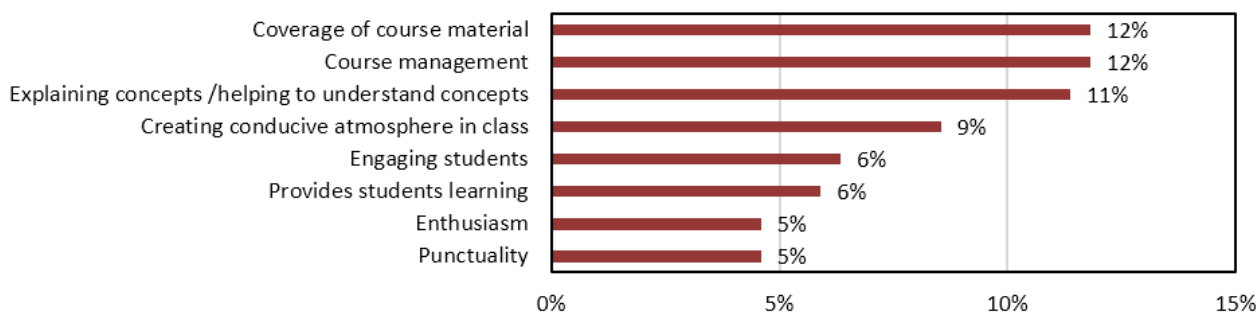

Figure 6: Factors of lecturer's effectiveness.

DISCUSSION AND CONCLUSIONS

The response rate for SECAT for the two semesters was low, as expected for a survey administered on-line. There was no correlation between the SECAT results and failure rate or average mark for the course. Despite the low rate of response it has brought some insight into students' concept of a good course and a good lecturer. Especially useful have been open-ended questions, which prompted students to report on their feelings and predilections. Students reported different preferences on course attributes for different types of course; but generally, students most liked it when a course was related to industry and practice.

A *model* lecturer explains concepts well, is enthusiastic, manages the class well, provides a conducive environment, is available and willing to assist, and is punctual. Whereas the signs of a *bad* lecturer are marking not fair and/or on time, bad delivery, no small group work, no hand-outs.

Students also indicated the most important attributes of lecturer's effectiveness was coverage of course material, course management, explaining concepts well and creating a conducive environment. Unfortunately, students also expressed their disappointment with the outcome of the SECAT exercise, which is repeated every semester. That disappointment is well summarised by the following student statement:

> *The truth is that I do not know why you always bother us with this SECAT form because we are still facing the same problems. There is no improvement to any problem that we state on this SECAT form. I have thought this was meant to improve leaning in this institution but it does not.*

Regrettably, the students' disappointment is well understood and genuine as SECAT is used only for staff assessment, not for teaching improvement or changes in the courses. It is seldom discussed. Students' assessment of teaching should be used as one of the ways to improve it. However, the evaluations *per se* do not induce change. Combined with self-evaluations they may focus academic staff attention on their own perception as instructors, and discrepancies between self and student evaluations may then motivate change. Such self-evaluation (self-assessment) should precede self-development. The recommendation is to perform the same SECAT assessment by academic staff, who can then check their own perceptions against students' perceptions and become aware of the discrepancies. Self-ratings that are mismatched with students' ratings are more likely to lead to changes in instructor behaviour.

REFERENCES

1. Guthrie, E.R., The Evaluation of Teaching: a Progress Report. Seattle: University of Washington (1954).
2. Brandenburg, G.C. and Remmers, H.H., The Purdue rating scale for instructors. *Educational Administration and Supervision*, 13, 399-406 (1927).
3. Zabaleta, F., The use and misuse of student evaluations of teaching. *Teaching in Higher Educ.*, 12, **1**, 55-76 (2007).
4. Davis, B.G., *Tools for Teaching*. San Francisco, CA: Jossey-Bass (2009).

5. Porter, S.R. and Whitcomb, M.E., Non-response in student surveys: the role of demographics, engagement, and personality. *Research in Higher Educ.*, 46, **2**, 127-152 (2005).
6. Centra, J.A. and Gaubatz, N.B., Is there gender bias in student evaluations of teaching? *J. of Higher Educ.*, 71, 17-33 (2000).
7. Boring, A., Ottoboni, K. and Stark, P., Student evaluations of teaching (mostly) do not measure teaching effectiveness. *Science Open Research* (2016).
8. Dillman, D.A., Eltinge, J.L., Groves, R.M. and Little, R.J.A., *Survey Nonresponse in Design, Data Collection, and Analysis.* In: Groves, R.M., Dillman, D.A., Eltinge, J.L. and Little, R.J.A. (Eds), Survey Nonresponse. New York: Wiley, 3-26 (2002).
9. Groves, R.M. and Couper, M.P., *Nonresponse in Household Interview Surveys*. New York: Wiley (1998).
10. Groves, R.M., Fowler, F.J., Couper, M.P., Lepkowski, J.M., Singer, E. and Tourangeau, R., *Survey Methodology*. Hoboken, NJ: Wiley (2004).
11. Groves, R.M., Survey Errors and Survey Costs. New York: Wiley (1989).
12. Porter, S.R., *Pros and Cons of Paper and Electronic Surveys*. In: Porter, S.R. (Ed), Overcoming Survey Research Problems: New Directions for institutional Research. San Francisco, CA: Jossey-Bass, 91-98 (2004).
13. Rubin, D.B. and Zanutto, E., *Using Matched Substitutes to Adjust for Nonignorable Nonresponse through Multiple Imputations*. In: Groves, R.M., Dillman, D.A., Eltinge, J.L. and Little, R.J.A., (Eds), Survey Nonresponse. New York: Wiley, 389-402 (2002).
14. Umbach, P.D., *Getting Back to the Basics of Survey Research*. In: Umbach, P.D. (Ed), Survey Research: Emerging Issues: New Directions for institutional Research. San Francisco, CA: Jossey-Bass, 91-100 (2005).
15. Groves, R.M. and Peytcheva, E., The impact of nonresponse rates on nonresponse bias: a meta-analysis. *Public Opinion Quarterly*, 72, **2**, 167-189 (2008).
16. Porter, S.R. and Umbach, P.D. Student survey response rates across institutions: why do they vary? *Research in Higher Educ.*, 47, **2**, 229-247 (2006).
17. Lindner, J.R., Murphy, T.H. and Briers, G.H., Handling nonresponse in social science research. *J. of Agricultural Educ.*, 42, **4**, 43-53 (2001).

BIOGRAPHY

Jacek Uziak is a Professor in the Department of Mechanical Engineering at the University of Botswana. He received his MSc in mechanical engineering from the AGH University of Science and Technology in Kraków, Poland and his PhD in technical sciences from the University of Life Sciences in Lublin, Poland. For the past 35 years he has been working at universities mainly in Poland and Botswana; his career includes teaching and research assignments also in Canada, Czech Republic, Norway, the UK, Netherlands, France, Germany and the USA. He specialises in engineering mechanics and teaches courses in this area. He has a particular interest in engineering education.