USING STUDENT FEEDBACK TO IMPROVE TEACHING

A Reflection

Junhu Wang

School of Information and Communication Technology, Griffith University,
Parklands Drive, Gold Coast, Queensland 4222, Australia
J.Wang@griffith.edu.au

Keywords: Student Evaluation of Course (SEC), Student Evaluation of Teaching (SET), Quality of Teaching, Performance.

Abstract: The use of SEC and SET in tertiary institutions to assess academic performance is widespread, and mandatory in some Australian universities. However, a lot of controversy exists about the practice. In this position paper, we argue that SEC and SET should be used mainly as a means to solicit constructive feedback for the improvement of course and teaching. We believe the current SEC and SET surveys in some Australian universities should be better designed, more carefully administered, interpreted, and used with caution. In particular, over-stressing interestingness of lectures and ignoring the quality of learning outcomes in the surveys should be re-considered in survey design. We propose strategies to help maximize the positive effects of SEC and SET on good teaching.

1 INTRODUCTION

Many universities nowadays use on-line student evaluation of course (SEC) and student evaluation of teaching (SET) to assess student satisfaction and presumably, the quality of courses and performance of lecturers. Some universities mandate that all courses and their teaching must be evaluated by students, and they use such evaluations as one of the criteria, and perhaps the only quantitative, claimed-to-be “objective” criterion in teaching, for academic performance review and promotion. Yet a lot of controversy exists with regard to this practice, see (Huemer, 2007) for a concise survey. Despite many years of extensive research and thousands of publications on it, the debate appears to be continuing among educationists, university management, and lecturers about the reliability and validity of SEC and SET, see (Davis et al., 2005; Reis and Klotz, 2011; Davis et al., 2007; Merritt, 2007; Centra, 2005; Cashin, 1990; Felton and Stinson, 2004; Grimes et al., 2004; Feldman, 1984) for some of the claims and counter claims. Supporters claim that SEC and SET are reliable, valid, and they are valuable means to help improving courses and teaching; Opponents claim such practice encourages indecent practice in teaching and the loss of academic integrity. In this position paper, we will review some of the major factors that contribute to SEC and SET, and present our own findings based on our personal experience as a teacher of IT for many years in Australia, as well as interviews with students and colleagues. We will argue that evaluation scores in a simplistic approach, without consideration of the extraneous factors, can be very misleading, particularly when the questions are poorly designed. If not used properly (e.g., when overused in academic promotion), it may create an environment that is counter-productive to good teaching. We propose strategies in the design, administration, interpretation and use of such evaluations that help to address some of the concerns caused by SEC and SET.

2 PURPOSE, RELIABILITY, VALIDITY AND USEFULNESS OF SEC AND SET

2.1 Purpose

For many university administrators, a main purpose of using SEC and SET is to assess the quality of a course and the effectiveness of a lecturer’s teaching, as perceived by students. Although it is common knowledge that students’ opinions may not reflect the truth, university administrators often treat students’ evalua-
tions as the predominant measurement of quality (perhaps because of the lack of cheap alternatives). In some faculties of Griffith University, lecturers who receive high evaluations are given certificates of appreciation and congratulation letters by the Dean. This is just an indication how highly student opinion polls are regarded by the academic managers.

A second purpose of using SEC and SET, which we believe should be a more important purpose, is to use it as a form of feedback from students, which will enable lecturers to take right actions to improve the course and teaching. Few people ever doubted the importance of student feedback in this respect.

A third purpose for some university administrators, we believe, and whether they admit it or not, is to gauge student satisfaction with the course, the teaching and the university. Many universities nowadays regard themselves as service providers (this is particularly true in Australia) and regard their students as customers. “Customers are the god” as the saying goes. And student satisfaction affects student retention rate, which is an important gauge used to assess the schools by senior managers, and even by various levels of governments.

2.2 Reliability, Validity and Effects on Quality of Teaching

Huemer (Huemer, 2007) defines reliability and validity as follows: A test is reliable if it tends to give the same result when repeated. A test is valid if it measures something it intended to measure. According to this definition, most people would agree that SEC and SET are highly valid in assessing students’ satisfaction, and research has also confirmed they are quite reliable as well (Centra, 1993). The main controversy is how reliable and valid such surveys are in assessing the quality of course and the quality of teaching, and whether the use of SEC and SET benefits good teaching.

Our main argument in this section is that the validity of SEC and SET depends to a large extent on how well the survey questions are designed, as well as how data are collected and interpreted; and whether the use of SEC and SET will benefit quality teaching is highly dependent on where and how they are used.

Certainly, the quality of course and teaching will have an effect on students’ evaluation, provided the questionnaires are properly designed, and there are students who are honest, bias-less, genuinely concerned with helping the university and lecturers make improvement. However, there are many other, sometimes more important, factors that may affect students’ evaluation, to such an extent that courses (or the teaching) with higher ratings do not necessarily have higher quality. Next we review some of these factors.

It was reported in (Marsh, 1987) that, when faculty members at a major university were asked what would most likely affect a student’s evaluation, 72% said course difficulty, 68% said grading leniency, and 60% reported course workload. Both our personal observation and interviews with students only further substantiate the views of these faculty members. In our own interview with 11 colleagues, all agreed that grading leniency, “kindness” to students, and course difficulty are among the major factors that affect student evaluation. 16 out of the 19 students who responded in a survey also agreed that the marks they receive will at least somewhat affect their rating, and all of the 8 international students we talked with informally (at social activities, without informing them this is a study about students’ attitude towards SEC and SET) said the “niceness” of lecturers is a major consideration when it comes to SEC and SET. Psychologically, it is not hard to imagine students who were penalized for plagiarism would rate the lecturer poorly as a form of revenge. It is just human nature that if the lecturer rewards students with high marks, he or she will be rewarded with high ratings by the students. This thinking is natural not only among students, but also for academics. Imagine academics are asked to rate their supervisors after performance review. Those who received comments such as “excellent”, “outstanding” from the supervisor will probably rate him/her much higher than those who received comments such as “poor” and “unsatisfactory”.

Our observation also indicates that the class size, the cohort of students (with various ethnic, social, and cultural background), and the nature of the subject field all have significant effects on student evaluation. Overall, in our school, students in smaller classes tend to give higher ratings to the lecturer, and the Master’s students (the majority of them are from overseas) tend to give higher ratings than undergraduates (most of whom are local students). Students who study hard tend to rate the course and teaching higher than those who often miss classes. Some of these observations appear to contradict the findings in (Davis et al., 2005), which claims that international students tend to give lower evaluations and class size does not contribute to student evaluations in first-year and second-year courses.

We believe that the design of survey questions (namely questionnaires) can also have an important effect on student evaluation. This is evidenced by the significant change of SEC and SET scores, for the same courses taught by the same lecturers using the
same teaching and assessment materials in an Australian university, after the number of survey questions were reduced by half, and the things asked by some of the questions were changed.

Previous research has found a number of other contributing factors that can significantly affect the score of SEC and SET. Among them, class venue and facilities, teacher appearance, native language, academic title, reputation in research, personality, and even age and gender. See (Davis et al., 2005) for a classification of these factors.

Given the many bias and factors that affect student evaluation, if the survey questions were designed improperly, and the data were used without consideration of these factors, the SEC and SET results will be unable to reflect the true quality of course and teaching. Moreover, if such SECs and SETs are overused in academic confirmation and promotion, they have the danger to degrade to popularity contests, and would encourage the drop of academic standards and the compromise of academic integrity. In the long run, it could do more harm than good to the quality of course and teaching.

3 DESIGN OF SURVEY QUESTIONS

Despite the issues with SEC and SET, we believe that using SEC and SET can prompt lecturers to more carefully reflect on their teaching practice. They are also valuable in collecting helpful feedback for improving the quality of course and teaching. In this section, we investigate the design of questionnaires used in SEC and SET and its effect on the validity of student surveys.

Ideally, the questionnaires should be designed as such that they can reflect the true quality of course and teaching (as much as possible), and encourage good teaching practice, while still being easy to use. Before looking at the questionnaires, let us reflect on what constitutes good teaching. Good teaching, to most people, means logical organization of contents, clarity in explanation, good pace, helpful and timely feedback, patience with weak students and so on. Good teaching in university must also encourage deep learning and help produce high quality learning outcomes (Kember, 2006; Biggs and Tang, 2007).

How well are the current SEC and SET questions designed? Do they assess the true quality and encourage good teaching practice? In late 2011, we surveyed 17 SET and SEC questionnaires from some major Australian universities, and we found that many of them could have been better designed. Some example questionnaires are attached in the appendix for easy reference.

Firstly, many of the current SECs and SETs fail to address whether the teaching encourages deep learning (such as critical thinking) and whether the course has produced good learning outcomes, while they stressed whether the lecturer has made the lectures interesting. In our opinion, this is one of the sources that has caused unfairness and controversy. Intuitively, the word interesting can mean different things to different people, but to most people interesting is related to entertaining and fun. However, many natural science subjects, by nature, is no fun (to the majority of students), and deep learning in these subjects require enormous personal efforts, sacrifice of entertainment time, which will translate to uninteresting. On the other hand, some social science subjects are, by nature, quite different. One can imagine a teacher teaching legal studies might start the lectures with case studies that relate to the relevant laws, and a teacher teaching history might start with telling stories and the like, but for someone teaching abstract mathematics, it is much harder to make it interesting. This perhaps explains why lectures in social sciences usually receive higher ratings than courses in natural sciences (Cashin, 1990) and why junior staff teaching easy courses often receive better evaluations than experienced professors who teach something technically more challenging. Human psychology also indicates difficulty is negatively proportional to interestingness (Myers, 2006). If you ask people what is more interesting, a story or a complicated mathematical proof, most people (who are not mathematicians) would say the former. Moreover, for a mandatory course, even if students are not interested in it, they must pass it in order to get their degree. For those who are not interested in the course but enrolled in order to meet the degree requirements, the course is unlikely to be interesting. Also, if a student is penalized for plagiarism or receives low marks because of poor work in assessments, the course is almost certainly not interesting to him. Therefore, we think that over-stressing interestingness in student surveys can be a major source of unfairness and inaccuracy. We think such questions should be replaced with questions such as “Does the lecturer explain things clearly, using a variety of pedagogical tactics?” On the other hand, many questionnaires do not contain questions that assess the learning outcomes perceived by students.

Secondly, we found some of the questionnaires are too short (perhaps in a bid to make the survey easy) so that they cover only some aspects of good teaching but ignore other important aspects. Also some questions are vague which makes them difficult to be an-
It is helpful to do the following: (of course), we believe that when doing such surveys, contributing factors for students' rating of a lecturer search (Marsh and Roche, 1997). Given the many and SET. This has been recognized by previous re-

important roles in the reliability and validity of SEC as well as the interpretation of the data can also play The data collection process, statistical calculation, as 4 ADMINISTRATION,

integrated for academic misconduct, or whether he or she has rated every other teacher equally low, whether the student failed the course before, or and whether the student is serious in providing honest feedback. Given that the vast majority of SEC and SET are done on-line, we believe such statistics is technically easy and practical. If the management is keen and interested in finding the truth, this should be implemented.

Most institutions we surveyed, but not all, also use two textual questions similar to the following: 1. What aspects of this teacher’s approach to teaching best helped your learning? 2. What would you have liked this teacher to have done differently? We believe these questions are more effective in soliciting specific constructive feedback from students, and they should be used in all student evaluations.

4 ADMINISTRATION,

The data collection process, statistical calculation, as well as the interpretation of the data can also play important roles in the reliability and validity of SEC and SET. This has been recognized by previous research (Marsh and Roche, 1997). Given the many contributing factors for students’ rating of a lecturer (or course), we believe that when doing such surveys, it is helpful to do the following:

1. Explain clearly to the students that the main purpose of such surveys is to help the teacher and institution make improvement, rather than as a basis to judge the performance of teachers. Ask participants to declare that they will provide honest, unbiased opinion.

2. Collect relevant data about the participants of the survey. Of course, such data should not be released to the teacher, but they can be used by the management to judge whether extraneous factors have affected the students' evaluation. For example, if a student’s rating of a teacher or course is extremely low, whether the student has been penalized for academic misconduct, or whether he or she has rated every other teacher equally low, whether the student failed the course before, or and whether the student is serious in providing honest feedback. Given that the vast majority of SEC and SET are done on-line, we believe such statistics is technically easy and practical. If the management is keen and interested in finding the truth, this should be implemented.

3. Analyze the evaluation result together with information about the course and class. Is the course a mandatory course? is it a first-year course? is it a difficult course? is it taught in a large class? is it taught first time? are the students mostly international students? How many students enrolled there failed the course before? We believe computer programs can be made that can put an appropriate weight to each of these factors and produce a “moderated” score of the evaluation.

4. Encourage more textual response. In particular, ask participants to provide textual explanation if he/she rates a teacher extremely low or extremely high. Give more weights to textual responses that use more objective, moderate language, and that provide more specific facts and suggestions.

5. Allow the teacher to identify students with whom there is a conflict of interest, and prohibit such students in the survey. In this way, the teachers will be less hesitant when it comes to dealing with academic misconduct.

6. Combine student evaluation with peer evaluation by colleagues, perhaps invited experts from other institutions. Peer-reviews are particularly important to assess the appropriateness of course content, difficulty of assessments, grading standard, and knowledge of the teachers. The peers should be experts in the subject field who have no personal relationship with the teacher.

Institutions should also establish a guideline as to how the survey data will be used. In particular, whether its teachers should seek to maximize SEC and SET scores. We strongly believe that SEC and SET should be used as a means to collect genuine feedback which can be used to improve the quality of course and teaching. It should not be overused in performance review and promotion, or even used pur-

\[\text{http://studentaffairs.stanford.edu/Registrar/faculty/courseevalspolicies}\]
\[\text{http://www.rucharacter.org/file/Microsoft\%20Word\%20-%20Teacher\%20Evaluation\%20by\%20Students(1).pdf}\]
posedly by some academic managers as an excuse to punish someone they do not like. In other words, lecturers should not be threatened by SEC and SET, so that they do not need to resort to inappropriate measures (such as flattering students with unfairly high grades) in their teaching.

5 SUMMARY AND CONCLUSIONS

In this position paper, we revisited the practice of using SECs and SETs in tertiary institutions, and provided some first-hand information based on our own observation, analysis, as well as surveys with students and colleagues. We believe the current SEC and SET questionnaires in some Australian universities should and could be better designed, more carefully administered, interpreted, and analyzed together with other information. We argue that SEC and SET should be used mainly as a means to solicit constructive feedback for improving the quality of course and teaching. If over-used in performance review and promotion, they can be counter-productive and negatively affect the quality of teaching. We proposed strategies to address some of the issues and help make SEC and SET a more positive thing on good teaching.

REFERENCES


APPENDIX

Questionnaire 1: Choose from strongly agree, agree, neutral, disagree and strong disagree. For the last question, choose from a scale of 1 to 7.

1. This staff member presented material in a clearly organized way.
2. This staff member presented the material in an interesting way.
3. This staff member treated students with respect.
4. This staff member showed a good knowledge of the subject matter.
5. Overall I am satisfied with the teaching of this staff member.
6. Overall, how effective was this lecturer/tutor in helping you to learn?

Questionnaire 2: Choose from a scale of 1 to 7, where 1=unacceptable, 7 = excellent.

1. How effective was this lecturer in helping you to understand what you were expected to learn?
2. How effective was this lecturer in using teaching methods that helped you to learn?
3. How effective was this lecturer in motivating and inspiring you to learn?
4. How effective was this lecturer in showing concern for you and your learning?
5. How effective was this lecturer in ensuring you received feedback which helped you to learn?
6. How effective was this lecturer in helping you to understand the standards of work?
7. How effective was this lecturer in helping you to extend your knowledge understanding and skills?
8. How effective was this lecturer in teaching in an organized, coherent, and well ordered way?
9. How effective was this lecturer in seeking and using feedback to improve his/her teaching?
10. Overall, how effective was this lecturer in helping you to learn?

**Questionnaire 3:** Choose from a scale of 1 to 5, where 1=strongly agree, 5 = strongly disagree.

1. The lecturer produced classes that were well organized.
2. The lecturer presented material in an interesting way.
3. The lecturer gave adequate feedback on my work.
4. The lecturer treated students with respect.
5. The lecturer seemed to know the course well.
6. The lecturer communicated his/her enthusiasm for the course.
7. The lecturer emphasized thinking rather than just memorizing.
8. The lecturer gave explanations that were clear.
9. The lecturer was available for consultation.
10. The lecturer helped me to improve my learning skills.
11. All things considered, how would you rate this staff member’s overall effectiveness as a university teacher?

**Questionnaire 4:** Choose from a scale of 1 to 5, where 1=strongly agree, 5 = strongly disagree.

1. The teacher let students know what was expected of them.
2. The teacher was well prepared for each class.
3. The teacher was accepting and supporting of students.
4. The teacher covered the material at an appropriate pace.
5. The teacher showed patience when dealing with students.
6. The teacher helped make the material interesting by relating it to things I know.
7. The teacher used visuals such as whiteboard, maps, handouts, and presentation stations.
8. The teacher handled student discipline fairly.
9. The teacher recognized which students did not understand the material and reviewed as needed.
10. The teacher appeared to be genuinely concerned about students and their success in class.
11. The teacher gave clear directions.
12. The teacher knew the subject matter well.
13. The teacher presented a professional appearance and behavior.

**Questionnaire 5:** Choose from a scale of 1 to 5, where 1=strongly agree, 5 = strongly disagree.

1. The staff member made the aims and objectives of the course clear from the outset.
2. The staff member made the subject interesting.
3. The staff member motivated me to do my best work.
4. The staff member provided adequate opportunities for me to pursue my own learning.
5. The staff member helped me to develop my understanding of concepts and principles.
6. The staff member displayed a genuine interest in my learning needs and progress.
7. The staff member gave me helpful feedback on how I was going.
8. The staff member used up-to-date teaching and learning approaches.
9. The staff member made it clear how her/his teaching developed the qualities of a University of XY graduate.
10. Overall, I was satisfied with the performance of this staff member.