The educational integrity of unit satisfaction surveys

Robert Kennelly
University of Canberra
rmkennelly@homemail.com.au

This work raises questions of educational integrity regarding the use of unit satisfaction surveys (USS) in an Australian university. The paper considers the educational integrity of using USS data through the prism of collaborative reflective practice programs designed to provide academics the space and safety to consider how they may contribute to improvements in the student learning environment. The paper does not make a judgment about the motivations for the introduction of a compulsory institutional, online unit survey, but on the impact on teaching and learning. It uses the experience of the author's co-facilitation of collaborative reflective programs for academics who wish to improve their teaching and their students' learning. It briefly considers the evaluation system that was replaced by USS and an alternative in another university. Finally the paper asks two questions of discussants: What does a unit satisfaction survey actually measure? And how do we encourage universities to implement student evaluation with educational integrity?

Keywords: student satisfaction surveys, improve student learning, educational integrity

Introduction

Context and rationale for work

The author is a co-facilitator of a series of *Talking About Teaching And Learning* (TATAL) programs run across universities in south eastern Australia. These programs seek to provide a safe environment where academics can reflect collaboratively on their students' learning and their own teaching (McCormack & Kennelly, 2011). It is in this context where the author encountered fellow academics reflecting on the findings of the unit satisfaction surveys (USS). Most TATAL participants have a Graduate Certificate in Higher Education and are well aware of the steps that are needed to improve student learning: understanding the complexity of student learning as outlined by Biggs in the 3P model (2003 p.19); and the importance of collecting timely, valid, representative data before embarking on the interpretation of such data which may lead to changes in student learning. The author became curious when he realised that a significant investment in improving teaching and learning appeared to be jeopardised by a compulsory online survey which offered considerably less to academics who genuinely wanted to improve their students' learning. This survey collects data on questions over which academics may not have control; is unrepresentative and untimely in its collection. Within the TATAL program it is counter productive for participants to be required to use the results of USS when they provide such questionable data on which to reflect on their students' learning (Kennelly 2011).

What is the USS?

- The USS is an online survey (see Appendix 1) that adopts some of the course experience questionnaire (CEQ) questions to establish whether students are satisfied with a particular unit. It is part of a broader framework described as The Course and Teaching Evaluation and Improvement system.
- It is available to students "from the beginning of the last two weeks of a teaching period and remains open until two weeks after results have been released." (University website).
- The results are used to improve teaching and learning and "are also available for use as part of professional development and review. There is no minimum response rate required for use of survey results" (University website).

- "Detailed results and benchmarking are distributed to the unit convener, Head of Discipline, Associate Dean (Education) and Dean" (University website).

What is educational integrity?

According to the Asia Pacific Forum Educational Integrity website 'educational integrity' is defined as "a commitment to the key values of honesty, trust, equity, respect and responsibility, and the translation of these values into action" (APFEI website accessed 25/06/11)

In the context of the *hypothesis* of this paper, educational integrity is that any considered enquiry into teaching and learning requires an understanding of the complexity of the environment in which student learning occurs; and a discipline to design and collect timely, representative and valid data which might allow for a considered interpretation and evaluation of the learning situation. Furthermore that the collection and dissemination of USS data is generally counter productive to those ends.

Literature research

There are three main points:

1. 'Satisfaction' is not an appropriate measure if the goal is to improve student learning. It is questionable as to whether satisfaction is a desirable outcome of university education (Richardson 2005). The popular notion of the student being the customer of the university needs to be considered against the background of what is to be learnt. Ramsden (2003, p. 221) argues that the fundamental question for students is "... about how effectively teaching engages them with learning...." Does the Australian taxpayer pay for happy customers or for graduate attributes? The measurement of graduate attributes goes well beyond whether the student was satisfied. Shelvin et al (2000) further muddies the waters when their study failed to identify or clarify the relationship between satisfaction and teaching ability or a student's personal view of a teacher. Finally Denson (2010, p.339) finds that "... Faculty selected optional questions are stronger predictors of overall satisfaction than compulsory questions". This would appear to run counter to the institution wide approach of USS.

2. The USS is not a reliable instrument for gathering student evaluation. Whilst the CEQ is based on sound psychometric testing it does not translate to unit level evaluation (Barrie & Ginns 2007). The notion that students are best able to provide all of the data required for the improvement of student learning is contested. Seiler and Seiler (2002) ask whether students possess the skills to know how a teacher should teach. Furthermore Nowell (2010) stresses the importance of testing the factors over which teachers have control. For instance a unit convener may not be in a position to pay casual staff sufficient money to allow for formative feedback to students on assignments. If we take q.7 of the survey (see Appendix 1) on feedback as an example and the unit convener does not have the funds to employ his/her tutors to provide feedback required by students then the unit convener is being judged on a matter over which he/she does not have control.

3. Improving student learning requires data from several different sources and considerable skill in interpreting that data. The evaluation of student learning is complex and is about individual differences (Prosser, 2005). Ramsden (2003) shares that evaluation requires the collection of evidence from several sources. Barrie and Ginns (2007 p.285) in their own work at Sydney University, suggest that evaluative data needs to be triangulated - data would include, valid representative student feedback, before and after learning including open ended questions, relevant assessment data and course evaluation focussing specifically at learning objectives. This would potentially provide a starting point for focus groups which would home in on potential areas for improvement in student learning. Prosser (2005) stresses the point that the key and overlooked part of student evaluation is the interpretation of the data once collected.

Advantages of USS

Some advantages associated with this survey include:

- Provides all students – should they wish to use it- with the opportunity to provide feedback on all units that they study. This includes open ended questions.
- Provides an institution-wide uniform source of data for student satisfaction of units and their teaching. Prior to the introduction of USS, CELTS surveys were used at the discretion of academics and in some cases not at all.
- Provides uniform data on students' perception of teacher performance.
- Good survey results are used to support the continuance of a particular unit.
- Some elements of the survey reproduce CEQ questions; CEQ results provide university funding.
- Opportunity to identify poorly performing teachers.
- USS focuses on the student experience.
- Any feedback is good feedback, especially qualitative comments.
- Online student evaluation provides more detail and is less costly than hard copy evaluations (Venette et al 2010).
- The use of USS and its role in the PDR is not static; improvements are being suggested; for instance closing the survey prior to the release of results at the end of the semester. Management sees USS as fitting into the broader framework of The Course and Teaching Evaluation and Improvement system.

Another advantage of student evaluation is its correlation with student assessment. Murray (2008) in his work in the University of Western Ontario found that students who perceive they are learning or being well taught also do well in unit assessment. However, Murray also points out that the results of student evaluations with expert development consultation produces significant improvement in teaching. At the university in question the availability of this expert consultation to all staff using USS is limited.

Disadvantages of USS

There are a number of criticisms of the surveys:

- The USS is a hybrid survey of 19 questions (University website) which are derived from the well researched instrument called CEQ which is designed to measure course satisfaction not unit satisfaction.
- USS questions are not teacher or situation specific.
- USS does not account for factors influencing student survey responses such as for rating bias between 'harder' and 'less difficult' units, bigger or smaller class sizes (Pounder, 2007), leniency in marking (Denson 2010 and Tang 1999), the tendency to punish teachers for low grades (Crumbley et al 2001) or the expectations of school leavers about so called 'hard' and 'easy' units, especially in science. (Kennelly 2011)
- Given the collection system the samples maybe unrepresentative (Bryant and Zhang 2010 and Denson 2010) and potentially skewed by timing the collection at the students' most stressed time.
- Some staff have criticised the way the USS is used as a 'stick' rather than a 'carrot' in the administering of the PDR process. (Kennelly 2011).
- USS does not have a minimum response rate so the same weight is given to a response rate of less than 20% as is given to one of 80%. (University website).
- A number of factors contribute to the ratings produced by the USS. A teacher's performance is only one of these factors (Nowell et al 2010).
- Most significantly, if the surveys were representative and valid they would not alone, be appropriate for measuring satisfaction or student learning as it would only be one aspect of a complex set of data required to reach a coherent view.

At the end of the day no amount of surveys, quality reviews, PDRs or evaluations will ever improve student learning (Ramsden 2003). It is in teachers conception of the complexity and individual differences that student learning can be understood and possibly improved.

Discussion

Have the USS designers assumed that the USS will deliver improvements in teaching and learning which will increase the CEQ satisfaction scores and provide the university with more funding? A reasonable assumption surely! However only at Kirkpatrick's (1994) first level of evaluation, *reaction* (smiley sheets), can this assumption be considered. That is USS data is asking students about their reaction to their unit experiences. At the second level, of a four tiered approach to evaluation, *learning,* the assumption is lost. Learning requires a before and after survey (Ramsden, 2003, p228) and consideration of the objective assessments of learning. An even more serious is the notion that students are in a position to give an objective assessment of their learning in a number of units whilst they are trying to submit their final assignment, prepare for multiple exams and come to grips with their results. Further, that the USS results will be used for PDR purposes regardless of the response rate or whether there is a representative sample.

The USS is designed to be part of the Course and Teaching Evaluation and Improvement system, a framework which includes a number of measuring strategies which aim to improve teaching. The author's experience with TATAL has generally not exposed the use of these other strategies. TATAL colleagues often appear singly focussed on the USS results and how to improve them.

Despite these criticisms the real tragedy for the author is that well motivated academics who are interested in improvements in their students' learning are being side tracked into using data from an inappropriate survey. Worse still they are in competition with those who mark leniently, have small classes and easier units. It is difficult to think of a more negative contribution to the challenge of understanding student learning.

Alternatives and what might have been

Barrie and Ginns (2007) argue that Sydney University have designed and tested a "unit of study evaluation system" (USE) which is consistent with their use of CEQ data. Even here Barrie and Ginns call on a triangulation of data to ensure a too narrow a focus is avoided (p.285). A sentiment echoed by Prosser (2005) and Murray (2008) who warn against relying on only student feedback which gives us only part of the whole story. Prosser goes on to say "So interpreting the results of student evaluation questionnaires as ratings for satisfaction are unlikely to result in major improvements for students" (Prosser, 2005 p.2) and places unnecessary pressure on staff to improve grades. We should be using the results to better understand the experience students have and homing in on particular problems (Prosser, 2005). Murray, 2008, shares the point about what student evaluations cannot pick up. Students can only provide feedback on what they observe...that is mainly in the class room. They cannot see the curriculum design, the planning of a lecture, the integrations of assessments skill development and the limitation of resources. These are parts of the experience which need to be examined by peer review, unit assessment or preparation for another more advanced unit.

The university had in place a highly reputable, sought after, and exported pen and paper face to face, anonymous feedback system which if used correctly took advantage of Kirkpatrick's level 2 evaluation, could be used in a timely way by the teacher concerned and provided a representative sample of the students in the unit under question. On its own of course this was insufficient data to use for analysis of student learning, but it provided valid student feedback. A compulsory version of this system is still being used at Trinity College in Dublin (Murray, 2011).

In summary the paper has considered the USS against the background of improving student learning; against the complexity of the environment in which student learning occurs; and against a discipline to design and collect timely, representative and valid data which might allow for a considered interpretation and evaluation of the learning situation. USS data can be used regardless of the representativeness of the sample. USS does not necessarily provide specific feedback for teachers. It measures 'satisfaction' which may or may not link improved student learning to a particular teacher. It considers factors over which the teacher may have little or no control. It provides a competitive environment where all teachers are not equal.

The USS appears to fail the educational integrity test in that it is not equitable and nor is it respectful of the development of individual teachers. Its translation into action is counter productive to a considered enquiry into teaching and learning and in interpreting a range of complex data which may contribute to the improvement of student learning.

Discussion questions:

1. What does a unit satisfaction survey actually measure?
2. How do we encourage universities to implement student evaluation with educational integrity?

References

Asia Pacific Forum on Educational Integrity website. http://www.apfei.edu.au/

Barrie, S. (2001). Reflections on Student Evaluation of Teaching: Alignment and congruence in a changing context. In E. Santhanam (Ed.), *Student Feedback on Teaching: Reflections and Projections*. Refereed Proceedings of Teaching Evaluation Forum 28-29 August 2000, Perth, Western Australia (pp. 1-16). Perth: Evaluation of Teaching Unit, Organisational and Staff Development Services, University of Western Australia. [verified 2 Sep 20911] http://www.osds.uwa.edu.au/other_services/publications/student_feedback_on_teaching

Barrie, S. & Ginns, P. (2007). The linking of national teaching performance indicators to improvements in teaching and learning in classrooms. *Quality in Higher Education*, 13(3), 275-286. http://dx.doi.org/10.1080/13538320701800175

Bryant, D. & Zhang, F. (2010). Uncovering unit level performance for first year students. *Proceedings First Year in Higher Education Conference*, Adelaide. http://www.fyhe.com.au/past_papers/papers10/content/pdf/11B.pdf

Chalmers, D. (2011). Student feedback in the Australian national and university context. In C. S. Nair & P. Mertova (Eds.) (2011), *Student feedback: The cornerstone to an effective quality assurance system in higher education* (pp. 81-97). Oxford: Chandos Publishing.

Crumbley, L. Henry, B. & Kratchman, S. (2001). Students' perceptions of the evaluation of college teaching. *Quality Assurance in Education*, 9(4), 197-207. http://dx.doi.org/10.1108/EUM0000000006158

Denson, N., Loveday, T. & Dalton, H. (2010). Student evaluation of courses: What predicts satisfaction? *Higher Education Research & Development,* 29(4), 339-356. http://dx.doi.org/10.1080/07294360903394466

El Hassan, K. (2009). Investigating substantive and consequential validity of student ratings of instruction. *Higher Education Research & Development,* 28(3), 319-333. http://dx.doi.org/10.1080/07294360902839917

Kennelly, R. (2011). Experience as a co-facilitator of collaborative reflective with teaching academics over the past 3 years [unpublished observations].

Kirkpatrick, D. L. (1994). *Evaluating training programs*. San Francisco: Berrett-Koehler Publishers, Inc.

McCormack, C. & Kennelly, R. (2011). 'We *must* get together and *really* talk…' Connection, engagement and safety sustain learning and teaching conversation communities. *Reflective Practice,* 12(4), 515-531. http://dx.doi.org/10.1080/14623943.2011.590342

Murray, A. (2011). Personal communications from the administrator of Trinity College student evaluation system.

Murray, H. G. (2008). Student evaluation of teaching: Is it valid and does it improve teaching. *Key note address, Australasian Higher Education Evaluation Forum (AHEEF) conference, Evidence based decision making: Scholarship and practice.* Monash University, October 2008. http://opq.monash.edu.au/cheq/aheef2008/conferenceproceedings.pdf

Nowell, C, Lewis, R. Handley, B. & Handley, G. (2010). Assessing faculty performance using student evaluations of teaching in an uncontrolled setting. *Assessment & Evaluation in Higher Education,* 35(4), 463-475. http://dx.doi.org/10.1080/02602930902862875

Pratt, D. D. (2005). Personal philosophies of teaching: A false promise? *Academe Online*, 91(1), Jan-Feb. http://www.aaup.org/AAUP/pubsres/academe/2005/JF/Feat/pratt.htm

Pounder, J. S. (2007). Is student evaluation of teaching worthwhile? An analytical framework for answering the question. *Quality Assurance in Education,* 15(2), 178-191. http://dx.doi.org/10.1108/09684880710748938

Prosser, M. (2005). Why we shouldn't use student surveys of teaching as satisfaction ratings. *The Higher Education Academy*, November 2005, 2pp. http://www.heacademy.ac.uk/resources/detail/ourwork /teachingandresearch/why_we_shouldnt_use_student_surveys_of_teaching_as_satisfaction_ratings

Ramsden, P. (2003). *Learning to teach in higher education* (pp. 209-232). London: Routledge.

Richardson, J. T. (2005). Instruments for obtaining student feedback: A review of the literature. *Assessment & Evaluation in Higher Education,* 30(4), 387-415. http://dx.doi.org/10.1080/02602930500099193

Seiler, V. L. & Seiler, M. J. (2002). Professors who make the grade. *Review of Business,* 23(2), 39-44.

Shelvin, M., Banyard, P., Davies, M. & Griffiths, M. (2000). The validity of student evaluation of teaching in higher education: Love me, love my lectures. *Assessment & Evaluation in Higher Education,* 25(4), 397-405. http://dx.doi.org/10.1080/713611436

Tang, S. (1999). Student evaluation of teachers: Effects of grading at college level. *Journal of Research and Development in Education,* 32(2), 83-88.

University of Canberra (undated). The Unit Satisfaction Survey (USS). [viewed 29 May 2011] http://www.canberra.edu.au/planning-quality/student-surveys-feedback/unit-satisfaction-survey

Venette, S., Sellnow, D. & Mcintyre, K. (2010). Charting new territory: Assessing the online frontier of student ratings of instruction. *Assessment & Evaluation in Higher Education,* 35(1), 101-115. http://dx.doi.org/10.1080/02602930802618336

Appendix 1: 2009 onwards USS Questions Scales

5 point Likert scale: Strongly Disagree, Disagree, Neutral, Agree, Strongly Agree. Students could also respond 'No opinion' with 'No opinion' responses treated as missing (no response).
% agree for scales with more than one question is the % of responses that were Agree or Strongly Agree. For scales with more than one question the definition used was that used by the Federal Govt 2009. Learning & Teaching Performance Fund (LTPF): to be considered a response the student had to have a response for at least 4 of the questions for the scale; the mean was then taken for these responses (with Strongly Disagree =1, Disagree =2, Neutral =3, Agree =4, Strongly Agree = 5); the mean was then rounded; rounded mean of 4 or 5 = Agree for the scale for the student.

USS: Unit satisfaction scale

q1: I found this unit intellectually stimulating
q2: The unit was well organised
q3: The unit helped me to develop skills and knowledge
q4: The methods of assessing student work were fair and appropriate
q5: I received feedback that assisted my learning

GTS: Good teaching scale

q6: The teaching staff put a lot of time into commenting on my work
q7: The teaching staff normally gave me helpful feedback on how I was going
q8: The teaching staff of this unit motivated me to do my best work
q9: My lecturers were extremely good at explaining things
q10: The teaching staff worked hard to make the unit interesting
q11: The teaching staff made a real effort to understand difficulties I might be having with my work

GSS: Generic skills scale

q12: The unit helped me develop my ability to work as a team member
q13: The unit sharpened my analytic skills
q14: The unit developed my problem_solving skills
q15: The unit improved my skills in written communication
q16: As a result of the unit I feel confident about tackling unfamiliar problems
q17: The unit helped me to develop the ability to plan my own work

OSS: Overall satisfaction scale

q18: Overall I was satisfied with the quality of this unit

SES: Student experience scale

q19: The unit made a positive contribution to my overall experience at the University of Canberra