

Deakin Research Online

Deakin University's institutional research repository

This is the published version (version of record) of:

Palermo, Josephine 2003, 20 years on - have student evaluations made a difference?, *in National quality in a global context ; Proceedings of the Australian Universities Quality Forum, 2003*, Australian Universities Quality Agency, Melbourne, Vic., pp. 136-140.

Available from Deakin Research Online:

<http://hdl.handle.net/10536/DRO/DU:30006226>

Reproduced with the kind permission of the copyright owner.

Copyright : 2003, Australian Universities Quality Agency

20 Years On—Have Student Evaluations Made a Difference?

Ms Josephine Palermo, Institutional Research Advisor, RMIT University

*Institutional Research Consultancy Unit, RMIT University,
GPO Box 2476V, Melbourne, Vic 3001, Australia*

This paper poses the question, what impact have student evaluations (SEs) of teaching had on the improvement in teaching in Australian tertiary education? The paper assesses the effectiveness of SEs through an investigation of the ways in which they have been used in Australian tertiary education over the last 20 years or so. Three approaches are discussed: a) quality assurance: SEs used to ensure the quality of the learning environment in which learning takes place, b) quality improvement: SEs used as a diagnostic tool for individual teaching staff, and c) student feedback: SEs used to provide prospective and current students with information to help inform their choices and expectations. The discussion reviews some of the research and known practice in Australian universities. It concludes that SEs may have had an impact on pedagogical understandings of tertiary education by placing the student at centre stage of teaching and learning. However, it suggests that we do not know how SEs have impacted on teaching effectiveness and improvement. The paper concludes by recommending further research to define critical success factors, a better understanding of how to make the student feedback and evaluation tools and mechanisms more approachable and meaningful for students, and an assessment of the 'survey industry' as a sustainable enterprise.

1. Introduction

In response to greater needs for accountability and improvement in tertiary education, many Australian universities have developed internal practices and procedures to collect and analyse student evaluations (SEs) of teaching effectiveness, particularly over the last 20 years (beginning with the seminal work of Herbert Marsh in 1984). This burgeoning activity, because it mainly relies on census collection rather than sampling techniques, has created a 'survey industry'. In most cases it is resource intensive and involves complex systems and of data capture and reporting. And to what end? An analysis of national mean scores from 1998 to 2001 shows that there have been no significant differences across any scales of the Course Experience Questionnaire (national survey of graduates administered by the Graduate Careers Council of Australia). Have Australian universities really capitalised on their significant investments in SE systems?

A way to assess the effectiveness of SEs is to understand the ways in which they have been used in Australian tertiary education over the last 20 years. Three approaches are discussed: a) *quality assurance*: SEs used to ensure the quality of the learning environment in which learning takes place, b) *quality improvement*: SEs used as a diagnostic tool for individual teaching staff, and c) *student feedback*: SEs used to provide prospective and current students with information to help inform their choices and expectations.

2. Student Evaluations—Investigating Their Uses and Misuses

It is generally agreed that 'good teaching' is context dependent on the nature of the learning goals and the quality of the environment in which learning takes place (Bowden & Marton, 1998, p. 130; Watson, Abril & Harvey, 2002). If this is the case, planned variations in learning experiences are the hallmark of good teaching. The *unplanned* variation in the quality of teaching and learning at RMIT was the second-biggest Top Ten issue in student feedback for 2002. Students are very concerned about the quality of teaching, and the impact of differing teaching methods on their learning experiences. The essential question for the University was 'how do we assure that learning is maximised in all given contexts?'

This means considering other factors that impinge on the learning environment beyond the instructor. Research from the Centre for Research into Quality at the University of Central England has shown that student feedback results on teaching performance are often confounded by perceptions about the adequacy of resources such as library and IT resources (Harvey, 2002). This would suggest that information required to improve good teaching should address students' perceptions of the quality of their learning environment, including learning support. Harvey (2002) has suggested that a generic set of questions about learning could be formulated around course organisation, the learning process, what students learn and learning support. He suggests that it should not incorporate questions about teaching performance. This approach, adopted by many Australian universities in 'unit' or subject evaluations, becomes the hallmark of quality assurance at the course and program level.

The debates on definitions of quality in education have led to ideas about 'accountability' and 'improvement' as dichotomised. This perceived dichotomy has been used to polarise approaches to the quality of teaching and learning in universities. Debates have focused around threats to academic freedom, and the dangers of accountability systems producing 'docile' and 'auditable' behaviours. More recently, there has been increasing support for the compatibility of accountability and improvement approaches (i.e. the Australian Universities Quality Agency and the New Zealand Academic Audit Unit). However, this tension is an important one in that the very nature of the SE system is dependent on how this tension is resolved.

3. Choosing Improvement as Assurance—Or Have We?

In many Australian universities this resolution has resulted in the use of SEs of teaching data being restricted to the determinations of the teacher/lecturer involved, and therefore in the main used for diagnostic purposes. For example, four of the universities in the Australian Technology Network¹ utilise system-wide mechanisms for evaluating student perceptions of teaching. In all cases, feedback information is used for improvement and implementation for action at the determination of the lecturer involved in the evaluation exercise. In this way data is often used to provide diagnostic feedback to staff about the effectiveness of their teaching. However, as Ballantyne, Borthwick and Packer (2000) note, there is little evidence that using such instruments in isolation has any effect on improving teaching performance. Instead, the research suggests that staff development interventions, including discussions with peers, consultants and follow-up support from 'expert' teachers, are more likely to impact on teaching improvement (Ballantyne et al., 2000).

The weight given to the use of SEs of teaching in institutions may be dependent on the importance assigned to teaching in the university. Read (2001) examined whether there was an association between (a) the relative emphasis placed on teaching in evaluating staff for tenure and promotion to full professor, and (b) the relative weight allocated to SEs in evaluating teaching performance. The results showed the relative emphasis given SEs was inversely associated with the weight assigned to teaching in staff evaluations for tenure and promotion to full professor. Findings suggested that, in general, program leaders who gave greater weight to teaching were likely to recognise SEs as providing only limited data on teaching performance. If these findings are generalisable, the conclusions they present are damning. They suggest that more research-intensive universities may use SEs more intensively because staff have less of a vested interest in teaching outcomes. Conversely, greater investment may require a more thoughtful and sensitive approach than SEs of teaching effectiveness.

4. And Whatever Happened to Student Feedback?

Prior to the 1990s, SEs in many universities were conducted by the students themselves, administered through student organisations, with results freely disseminated in 'alternative' course guides. It is interesting to note that, 20 years on, this practice has almost disappeared from the Australian tertiary landscape, usurped by systems that are mainstreamed and controlled by university administrations. The

¹ University of Technology Sydney, Queensland University of Technology, University of South Australia, Curtin University of Technology.

strengths of student organisations, particularly in their ability to encourage students to be effective partners in learning with university administrations/faculties, were substantially diminished through regressive legislation such as Voluntary Student Unionism. Therefore, their prominence in the debates on SEs and their uses has been obscured. So much so that today while students are recognised as 'clients' they are often not the recipients of SE results and meaningful feedback, but rather more likened to 'inputs' to complex SE databases and processes. The notion of student feedback seems to have been turned on its head.

Feeding back results to students in a meaningful and purposeful way remains a challenge for student survey administrators. Surveys are not always timely and conducive to feed back for students, and the type of information 'shared' is not always interpreted as meaningful by students. What is missing often is data that reflects the kinds of issues that are important to students. These often implicit narratives are more likely to be unearthed through qualitative data techniques, the importance of which has been recognised, for example, in a collaborative research project headed by researchers at the University of Technology Sydney and the Queensland University of Technology, on the use of qualitative responses to the Graduate Careers Council of Australia Course Experience Questionnaire. However, even then, issues of who 'owns' the data obstruct opportunities for feedback to students. One interesting trend to note is the proposal for a national 'alternative' course guide in the recently released United Kingdom White Paper on educational reform. Time will tell whether this may mean a move towards a more student-centred approach to SEs in the United Kingdom.

5. Do Student Evaluations Place the Student at the Centre of Teaching?

Having identified some of the issues surrounding the uses of SEs over the course of their proliferation in Australian universities, it is difficult to conclude that SEs overall have had an impact on teaching improvement. However, it can be said that SEs have made student perceptions of teaching more salient in the thinking around the scholarship of tertiary teaching. Does this equate to an improvement in teaching effectiveness? To better answer this question the discussion next turns to the reliability and validity of SEs and the conclusions we can draw from their use.

Research on the reliability of SE instruments is most often specific to a single instrument and has focused on three areas: consistency or inter-rater reliability, stability, and generalisability (Hobson & Talbot, 2001). While the debate about the reliability and validity of SEs has produced a plethora of empirical data, the research does seem to conclude that there is a degree of consistency overall in SEs given class sizes are greater than 15 (Cashin, 1988, as cited in Hobson & Talbot, 2001), that SEs are relatively stable over time and that results are generalisable across courses and students (Marsh & Overall, 1981). In addition, research suggests that students' overall ratings correlate positively with how much they learn, are not affected by their own personal characteristics or grade point averages, and are not influenced unduly by the instructor's personality or style (Marsh & Roche, 2000).

Validity of SEs is more difficult to confirm because researchers concede that there is no universally accepted criteria for what constitutes effective teaching (Feldman 1988; Marsh 1984). In addition, correlations between peer ratings and student ratings have been shown to be unacceptably low (Marsh, 1984). So while we can conclude that SEs are reliably measuring student perceptions, the nature of those perceptions, and how they relate to teaching effectiveness, remains somewhat obscure.

6. Closing the Loop on All This Data

Despite the known problems with SEs, what is clear is that simply collecting student feedback is not enough; there must be some purposeful actions that accompany the capturing and dissemination of information to those best placed to do something about it. Often the sharing of results from SEs is complex within the confines of industrial relations and institutional cultures that attribute differing degrees of 'blame' to staff. Tensions between management, staff and students are real and need to be negotiated before agreements on practices can be achieved. What is clear, however, is that such negotiation is possible.

For example, at the School of Physiotherapy at Curtin University (CEW), staff reached a collective agreement about the use and nature of feedback results of teaching and learning. Information was shared at all levels within the school in the spirit of improving learning for students. This example demonstrates the need to focus on issues at a local level that reflect what is that is important to students, and to the teaching staff involved.

Neumann (2000) suggests that in reporting rating results, it is important to consider that within the university community there are multiple evaluation users who have different purposes for which they need and use evaluation results. He identifies four types of users: (1) the individual academic requesting an evaluation of their course and teaching, to reflect on their own teaching for improvement purposes and to use results of their evaluations for a variety of career purposes; (2) academics who are members of selection, tenure and promotion committees; (3) staff developers who are called on to advise members of the university community on all aspects of evaluation, and implications for decisions about teaching quality and improvement; and (4) students at the very least who like to know that the time they have spent completing surveys brings about genuine change that will impact them directly.

The identification of users, and the subsequent broadening of stakeholders who have some legitimacy over the information elicited through student feedback, broadens the nature by which results are tied to action and improvement strategies. A review of the use of SEs at the Queensland University of Technology (Gordon, 1996, as cited in Ballantyne et.al., 2000) encouraged the use of these evaluative instruments to support developmental (formative) processes within the University. It recommended the integration of staff development and evaluation practices. By directly linking professional development and course improvement strategies to SE outcomes, staff can be encouraged to take follow-through action on information obtained through the use of SE instruments

A trawl through university web sites in Australia does suggest that similar strategies are linked to the uses of SEs. However, more evidence is required that these strategies and actions are implemented. More research is required that shows that these improvement activities are actually contributing to improvements in teaching and learning. The same perusal of web sites shows that while students and staff of the particular institution may be privy to findings and subsequent improvement plans and actions resulting from SEs, the public in general is often left out of the loop. This is important when we consider that stakeholders such as industry partners, prospective students and their parents, government and community agencies are represented in this group.

7. Concluding Remarks

This paper posed a question about the impact that 20 years of SEs of teaching has made on the Australian tertiary education system. While this discussion has by no means been exhaustive, there are some conclusions that can be made and some further questions that should be asked. From a sectoral perspective, we can presume to know that SEs may have had an impact on pedagogical understandings of tertiary education by placing the student at centre stage of teaching and learning. This shift to more student-centeredness may have occurred anyway, but the salience of student perceptions and evaluations is undeniable. However, we cannot presume to know how the proliferations of SEs have impacted on teaching effectiveness and improvement. We need further research to define critical success factors for teaching and learning improvement at an institutional level that include an evaluation of current practices surrounding data collection, analysis and reporting of results. These may not necessarily be quantitative indicators. We also need to better understand how to make the student feedback and evaluation tools and mechanisms more approachable and meaningful for students, and how to re-energise the 'student voice' around these issues. We need to assess whether the 'survey industry' is a sustainable enterprise, and where it may be headed in the next 20 years. And we need to assess how far we've come from the seminal work of Marsh (1984) and whether it's worth going further.

References

- Ballantyne, R., Borthwick, J., & Packer, J. (2000). Beyond student evaluation of teaching: Identifying and addressing academic staff development needs, *Assessment and Evaluation in Higher Education*, 25(3), 221–236.
- Bowden, J., & Marton F. (1998). *The university of learning*, London: Kogan Page.
- CEW, *Curtin Business School*. Retrieved from <http://physio.curtin.edu.au/cew/introduction.shtml>
- Feldman, K. A. (1988). Effective college teaching from the students' and faculty's view: Matched or mismatched priorities? *Research in Higher Education*, 28, 291–344.
- Harvey (2002). *Report on the Student Experience at UCE*, Birmingham: Centre for Research into Quality, University of Central England.
- Hobson, S. M., & Talbot D. M. (2001). Understanding student evaluations, *College Teaching*, 49(1), 26–30.
- Marsh, H. W. (1984). Students' evaluations of teaching: Dimensionality, validity, reliability, validity, potential biases, and utility, *Journal of Higher Education*, 46, 89–102.
- Marsh, H. W., & J. U. Overall. (1981). The relative influence of course level, course type, and instructor on students' evaluations of college teaching, *American Educational Research Journal*, 18, 103–12.
- Marsh, H. W., & Roche, L. A. (2000). Effects of grading leniency and low workload on students' evaluations of teaching: Popular myth, bias, validity, or innocent bystanders? *Journal of Educational Psychology*, 92(1), 202–28.
- Neumann, R. (2000). Student evaluation of teaching results: Rating interpretation guides, *Assessment and Evaluation in Higher Education*, 25(2), 121–134.
- Watson, S., Abril, S., & Harvey, L. (2002). *The 2002 report on the student experience at UCE*, Birmingham: Centre for Research into Quality, University of Central England.
- Read, W. J. (2001). The relationship between student evaluations of teaching and faculty evaluations, *Journal of Education for Business*, 76(4), 189–192.