Quality in teaching and learning

One path to improvement

Geoff Sharrock

L H Martin Institute, University of Melbourne

In response to the concerns raised by John Buchanan in his recent article about the uses and abuses of student surveys (AUR vol. 53, no. 1), this paper outlines one method for using student feedback to inquire into the quality of teaching and learning in classroom settings, and identify improvements to this and other aspects of a program of study at the subject level.

Introduction

I read John Buchanan's article just after receiving the results of a short 'quality of teaching' survey with nine items and a rating scale, widely used at my own institution, for a subject I had taught in 2010. I had given the survey to my students in their final class, in October. The completed forms were mailed off to be processed, and I received the summary results back in March, with an average score for each item on the list, and some supplementary comments where students had elected to offer these. As a relative newcomer to academia, and as someone new to this programme, teaching a subject with few pre-existing course materials, I was happy to see that, having scrambled to be ready for many of the classes, my student ratings weren't bad: there was room to improve, but no cause for concern.

Clearly this process serves reporting purposes of value to a university (particularly one such as mine which, having made big changes to its curriculum, can now present aggregated 'quality of teaching' survey data captured over several years which show that, despite what some critics may claim, the Melbourne Model has merit). However, while recognising these benefits, I can also say that at the front line of teaching, the information generated by this survey was not much help to me at all. It did serve

to reaffirm my own sense that all was indeed well, and to reassure my programme coordinator of this. But, as a systematic way of using survey-based student feedback to gain insights into the dynamics of a specific teaching and learning situation, and generate meaningful improvements in that situation, the approach taken here was too static, too anonymous, too slow and too simplistic. On the last point, for example, the content of the survey propositions underplayed the extent to which aspects of the learning environment, positive or negative, are co-produced by students themselves (for example, by doing the reading, turning up, asking questions, engaging in thoughtful and open discussion - or failing to do these things).

In practice, a quality of teaching survey designed to be short, sharp and simple to administer too easily becomes a blunt instrument, wielded slowly and indiscriminately, to yield data that are, diagnostically speaking, dead on arrival. Student survey processes like this seem designed to suit the needs of central planners, faculty managers and policy researchers far more than those of front line practitioners, raising the kinds of 'fitness for purpose' concerns that Buchanan highlights. No doubt the delay I experienced between seeking and getting student feedback can be minimised by adopting an online approach to administering surveys, as some institutions have done (and as my own insti-

tution now does). And a longer, more detailed list of survey propositions, with some optional extras, would yield a more fine-grained snapshot of students' experiences of a particular programme or subject (as my own institution allows for).

Typically, however, fixing one problem exacerbates another. Depending on how it is administered, a significantly longer survey may lower the response rate where participation is voluntary. As Buchanan notes, a survey administered in class is more likely to have a high response rate than one undertaken out of class and online. How then to ensure that the process will be systematic, transparent, user-friendly and also detailed and nuanced enough to be of use to practitioners seeking meaningful improvements from meaningful feedback? The conventional processes for capturing, consolidating and sifting through student feedback seem to offer little scope for a lecturer or programme coordinator to:

• engage students in a conversation that elicits the story behind their individual responses;

• assure students that their specific concerns are understood and taken seriously;

• assure students that where it makes sense to do so, changes can be made; or

• communicate to students why a proposed change is unnecessary, infeasible or unwise.

In sum, many of the methodological concerns that Buchanan and others have raised about the uses of standardised teaching quality, course experience or student satisfaction surveys to improve the quality of provision remain unresolved. The survey items may be clear, their relevance to student learning may be well supported by research, and the ratings produced may, in aggregate, be statistically valid. But this is small consolation for the fact that, from an improvement point of view, the information produced by these processes offers little of immediate value to the practitioner. The indicators produced are not very informative, and tend to be all 'lag' and no 'lead': if a problem needs attention, the survey data flagging this tends to land too late, and when it does land, offers few clues as to what exactly should be done differently, by whom and for whom. Where the prompt for change is expressed by a minority of students, the lecturer has no firm basis for knowing whether a particular change of approach will suit the majority of students in that class, let alone the next.

All that said, Buchanan's critique of the uses and abuses of student surveys does not go on to propose a more effective and appropriate alternative, which would enable the more constructive, formative uses of student feedback he seeks. Nor is it enough to object that a reliance on student surveys 'presumes that students know more about edu-

cational quality than do their teachers' and misguidedly privileges the 'student as client' (2011, p. 67). The more nuanced view offered by James, for example (2002, p. 79), seems more persuasive:

> Since student expectations must have some bearing on their motivation and satisfaction, expectations must in turn influence the quality of higher education for students are co-producers of this quality…Students are well-equipped to judge the quality of certain aspects of higher education…Students can be expected to be reasonable arbiters of the impact on them of the availability of computers, the quality of teaching spaces, the teaching skills of academic staff, and so on… They expect the fundamentals of effective teaching — clear goals, feedback on progress, and transparent assessment requirements and grading practices — and they welcome personal interaction with teaching staff and being treated as individuals by staff who show concern for their progress…But the student expectation-quality relationship is not altogether this straightforward…There are deeper dimensions to quality in higher education, such as the overall coherence of the curriculum, into which students have fewer insights. These aspects of higher education quality are usually less tangible, less intuitive and require a longer term view. Students are not necessarily in the best position to judge these aspects of quality, creating potential clashes between individual student preferences and what is educationally desirable…

Common indicators of 'quality' or 'excellence' in higher education are well represented by the Australian Teaching and Learning Council (ALTC) criteria which have 'become accepted as a proxy list of skills and practices of effective university teaching in Australian higher education' (Devlin & Samarawickrema, 2010, 115-118). The ALTC list highlights engaging and motivational teaching; coherent and up to date course material; both formative and summative assessment and feedback; support for individual student development; and reflective and evaluative practice. Also, several universities have excellent material available online to guide teaching and learning practitioners, solidly based in a wide range of scholarship. The University of New South Wales, for example, has a framework based around 16 principles (UNSW, 2004), and the University of Melbourne, has one based around nine principles (University of Melbourne, 2007). Both frameworks are widely referenced, well-articulated and user-friendly, translating robust principles of teaching and learning into detailed lists of good practice that practitioners can use to reflect on their own approach, and build their repertoire.

The student surveys I have seen are not inconsistent with any of this. But the basic design problem, which Buchanan's analysis illustrates so well, is that the systematic, standardising and summative aspects of conventional

Quality in teaching and learning, Geoff Sharrock **79**

student feedback processes tend to trump the local practitioner's need for more individualised, context-specific and formative processes - just as they do with other university management processes, such as quality assurance generally, or staff appraisal. Hence the difficult balancing act, where the design of student feedback processes fails to bridge different educational and administrative priorities, and formative and summative purposes. As Biggs observes (2001, pp. 231-232):

> Many institutions have mandatory SFQs [student feedback questionnaires] as summative evaluations at the end of each course, using standard questions across all courses, where the lecture is assumed to be the norm. Ratings then vary according to students' own conceptions of teaching, and penalise teachers using other methods… Used formatively, however, SFQs make eminent sense where questions are tailored to specific courses on aspects [where] the teacher wants feedback…

None of the observations I have made so far is new: I have no grand solutions to offer, that would resolve these issues at an institutional level. However, the rest of this paper will outline an approach to working with student feedback at the subject or programme level, which goes some way toward bridging the summative-formative gap. For the past few years I have been using this in programmes where I have played a coordinating or lecturing role.

Using an 'interactive survey' as a basis for shared inquiry

Figures 1-9 below present sample images of an 'interactive survey' approach I use to capture, graph, play back and discuss student feedback data in a live group setting. The aim of the exercise is to map, disclose and examine how students, individually and collectively, experience various aspects of the subject, and the approach taken to organising and teaching it. My toolkit for doing this is a one page survey with 42 propositions and a nine point rating scale; a Windows based software platform to capture and graph the data; and a laptop and data projector to present it.

For small groups of (say) 20 students, the process takes up to 60 minutes: 10 to fill out the forms, 15 to enter the data (usually in a short break), five to give the forms back, and 30-40 minutes for the group to explore the results, identify issues, diagnose causes, and consider possible improvements. As the summary results are played back, students are asked to compare notes and interpret their ratings of the survey items. Their main tasks are to:

- See how their individual ratings of items compared with those of others in the group.

- Volunteer explanations as to why they scored particular items as they did.
- Discuss whether low scoring items represent a significant problem for the whole group, a problem for a few individuals, a minor concern, or a misunderstanding.
- Volunteer suggestions as to how any serious concerns discussed might be dealt with.
- Consider whether any changes proposed are feasible and beneficial, and would not downgrade some other valued aspect of the subject.
- In some cases, adjust expectations that, on closer examination, seem unrealistic.

By the end of such a discussion we may have identified two or three issues that warrant attention, and (unless a problem is too intractable to resolve easily) agreed on what kind of change would offer benefits, if adopted. The immediate follow-up action is to report to the group with a snapshot of the most relevant parts of their survey data, with some brief notes that state the problem(s) identified and the solution(s) to be tried. Usually I can do this a day or two after the discussion takes place, while memories are still fresh, by copying a selection of screen shots into a Word document, and then emailing or posting the report. The follow-up action is to implement any changes agreed on, and see if these improve the learning situation.

How interactive survey data summaries are presented to students

Figure 1 gives an overview of the seven themes used to organise and present summary scores for my 42 survey propositions, with each theme representing a set of six propositions. The Overview shows the average score across the whole group for each theme, in this case for a class of 24 students at an early stage in the programme. Since all 42 survey propositions seek to represent 'good practice', higher scores are seen in principle as positive (9 = Always), and lower scores as negative (1 = Never). But this is not always so. In discussing the examples people had in mind when rating a particular item, I can invite
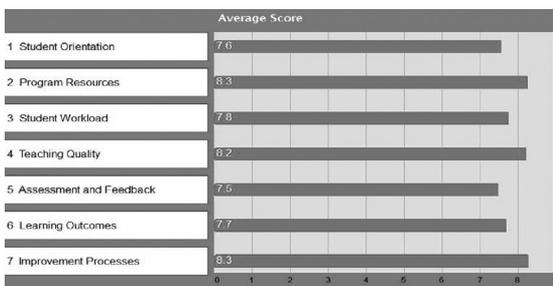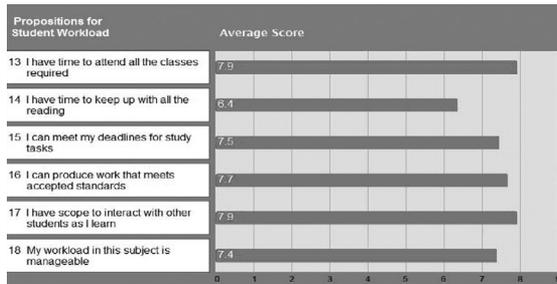

Figure 1: Overview

Figure 2: Student Workload

them to challenge the survey's assumptions, and suggest better propositions that would help to pinpoint the issue we are seeking to diagnose. This summary view gives a first impression of whether a group's overall experience has been positive; and if not, it indicates roughly where problems might be found. Usually there is no need to discuss the Overview summary beyond noting that the survey data is arranged according to these general themes: the discussion points come later, when considering and comparing data at a more detailed level.

Figure 2 gives an overview of propositions for one of the seven themes, Student Workload. This second level of data shows average scores across the group (in this case, 16 students) for each of the six items in the set. Here I ask students to note which items have low average scores. This is where the issues are likely to emerge, to be probed and interpreted.

Figure 3 shows the third level of data: how students' scores for each survey item are distributed across the group, as a percentage of all scores for that item. Typically, this screen highlights how diverse student experiences within a group can be. Here I ask students to check back to see if their scores are above, below or in line with the group. In this case, in a group of 11 students, two were always able to keep up with the reading, two were able to do so less than half the time, and the rest sat somewhere in between. Since nearly half the group had struggled with the reading load, we discussed this. I found that my reading list comprised a larger load than in their previous subject, which had set expectations. (Next time, listing core and non-core readings may help.)
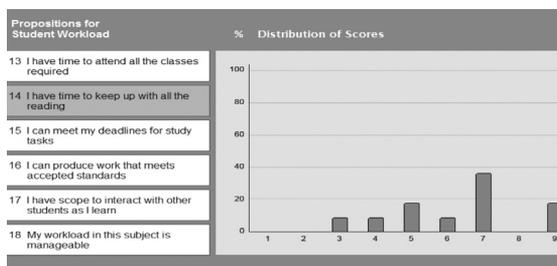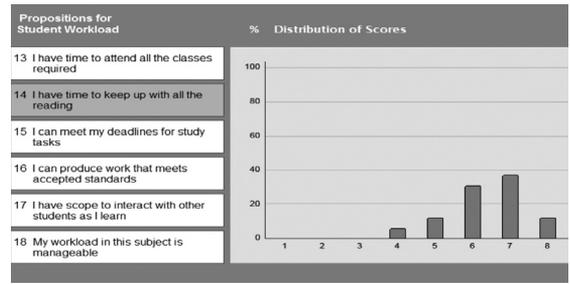

Figure 3: Student Workload

Figure 4: Student Workload

However, the volume of set reading may not be the issue. Clearly, a large part of students' experience of course quality is essentially administrative. Their learning experience will often depend on things like: the course materials being available well in advance, and in a clearly set out and accessible form; ready access to the physical venues where classes take place; having enough power points in the room to plug in laptops; and getting assistance with the technical aspects of online procedures. In Figure 4 for example, the discussion of workload associated with 'keeping up with all the reading' in a different class turned out not to be a concern with the volume of reading, as the students found it relevant and did not want less. The real concern for some was how soon they would have access to the material, to prepare for classes and study tasks. The context for this group of 16 students was that as busy managers, studying part-time, often they had only small windows of time for study, due to business trips and the like. In response the first batch of course materials in the next subject was made accessible online some weeks ahead of the complete set. (Normally this became available only after the complete set was revised, with updated course notes, after different contributors had all had their input, so that hard copies in folders could be mailed out).

Implications

These examples illustrate two simple points about the reliability of student survey data:

- as shown in Figures 3 and 4, different classes may give the same average score for an item for quite different reasons (in this case, 6.4 and 6.3 for the same proposition); and
- without a group discussion of survey results, the salience of the issue of timely access to course materials may not appear on the lecturer's radar at all, unless a survey item expressly seeks this information.

Regarding pedagogy, much depends on the preferred learning styles of students themselves. In the case of item 21 (Figure 5), I asked the lower-scoring respondents, who had rated my style of teaching as less engaging than it
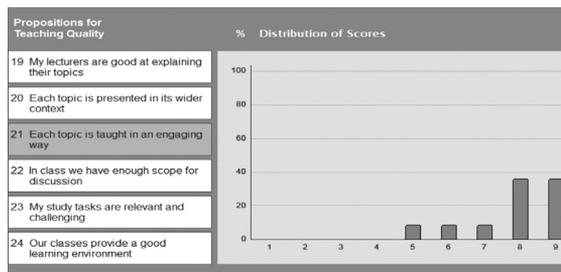
Propositions for Teaching Quality	% Distribution of Scores
19 My lecturers are good at explaining their topics	
20 Each topic is presented in its wider context	
21 Each topic is taught in an engaging way	
22 In class we have enough scope for discussion	
23 My study tasks are relevant and challenging	
24 Our classes provide a good learning environment	

Figure 5: Teaching Quality

Propositions for Improvement Processes	% Distribution of Scores
37 Our lecturers seek feedback from students	
38 Our lecturers take student feedback seriously	
39 If a problem is raised, it is given prompt attention	
40 If there is scope to improve, changes are made	
41 If student demands are unrealistic, this is clarified	
42 Our lecturers show commitment to improvement	

Figure 6: Improvement Processes

might have been, to volunteer specific comments and suggestions. I found that some would have liked me to widen the mix of activities, and rely less heavily on a 'show, tell and discuss with PowerPoint slides' routine. This I can do.

Discussion – limitations and benefits

This kind of process won't be for everyone, obviously. If things are going badly, it may be too hard for the lecturer to process the feedback alone in a live group setting. It relies on a fair bit of trust between the lecturer and the students, since students may otherwise fear that any criticism they make may cause offence and come back to haunt them later, when the lecturer assesses their work. And it has some technical limitations, particularly if used in larger classes, where the process of collecting and entering the data can consume too large a chunk of the time available to the class. (In the always-online, iPad world now emerging, a technical solution for this is likely to be one where students enter their scores directly, get individual profiles of results on their own screens, and view the consolidated results on a large screen). But for someone in my kind of situation, teaching fairly small groups of postgraduates in a fairly low-tech seminar format, where there is typically a fair bit of open discussion, and with well-motivated students who usually display the 'amicable co-operation' that Buchanan so rightly recommends (2011: 69), there are perhaps a dozen potential benefits:

- First, the process is inherently educative, with the survey results used as a qualitative instrument of active inquiry, not (just) as a passive, quantitative measurement.
- Second (unless class attendance is unusually low), a high response rate is guaranteed, leading to high confidence that the summative data is representative.
- Third, it is not hard for the lecturer to make the process of mapping and probing student feedback an engaging and inclusive one for students.
- Fourth, the process is extremely transparent, because the data are collated in front of respondents, and the results are available almost immediately.
- Fifth, the findings are highly reliable, since interpretations of the various ratings are confirmed, qualified or
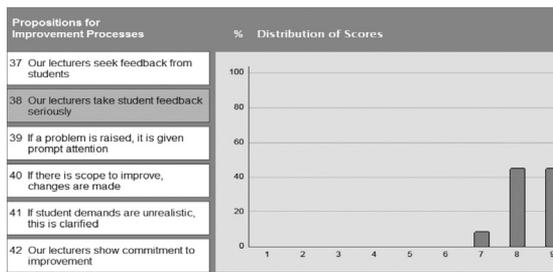
discounted in discussion with those who did the rating.

- Sixth, individual students learn whether their own concerns and judgements are widely shared, or are out of step with the experiences and perspectives of others.
- Seventh, there is scope to use a collective 'differential diagnosis' of the data to challenge individual students' assumptions or perceptions in a methodical, constructive way.
- Eighth, the feedback process tends to evolve rapidly from one-way feedback to the lecturer, to two-way dialogue between lecturer and student, to multi-party dialogue across the group, creating the conditions for mutual adjustment.
- Ninth, there is scope to test with the group whether a proposed change of approach seems likely to offer the kind of benefit sought, without collateral damage to other benefits of value to the group (or to wider constituencies), that flow from the existing approach.
- Tenth, there is scope to report quickly and easily in some detail to the students, and to colleagues or supervisors, on how students are experiencing a subject, what concerns they have raised, and what steps can sensibly be taken to improve their learning experience.
- Eleventh, having detailed and verified quantitative data at one's fingertips offers scope for a lecturer or department head to respond readily and credibly to any uninformed or unfair criticisms that may be circulated by disaffected students on Facebook, unsupportive colleagues in faculty networks, or overly competitive rivals in other institutions.
- Twelfth, the discipline inherent in the process, of gathering and handling student feedback responsibly and transparently, cuts both ways: the lecturer must take student views into consideration, but the views that will carry most weight are those that represent well-considered feedback – and ill-considered feedback is soon revealed as such.

Overall, the process allows reasonable scope to bridge the formative and summative purposes that student surveys are supposed to fulfil, while leaving room for practitioners to exercise professional discretion in their handling of student expectations and concerns.
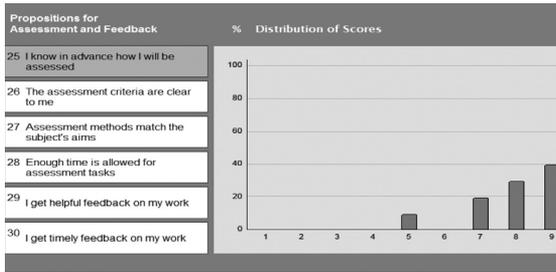
Figure 7: Assessment and Feedback

As Figure 6 indicates, the lecturer can demonstrate that student feedback is valued, without having to accept it uncritically. The lecturer can make judgements and give reasons as to:

- Which student concerns warrant a response as a matter of priority.
- Which should be recognised as relatively unimportant, given the context.
- Which should be recognised as valid concerns that may not be resolvable given other priorities or resource constraints, and
- Which reflect an unrealistic expectation.

Further, as the person controlling the technology, the lecturer can clarify the stories behind the data, unravel mixed messages, and test whether any particular viewpoint is as representative of the group's experience as its proponents claim. (Without fine-grained data at its fingertips, it is not hard for a group to slip into relying too much on the assertions, impressions or anecdotes offered by one or two vocal individuals). Because students review in a group setting how they rated their individual experiences of the subject, the process tends to help individuals moderate their own sense of which aspects or the current situation are wholly positive or wholly negative, and how much weight to give these. This takes pressure off the lecturer to defend and justify his or her approach in the face of any strident or dogmatic claims about how things should be done, why something is not working, or what the lecturer should do. This applies particularly where a student's negative view of some element of their experience is due, at least in part, to a misconception, or their own inattention.
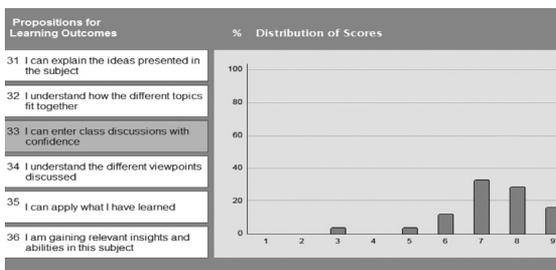

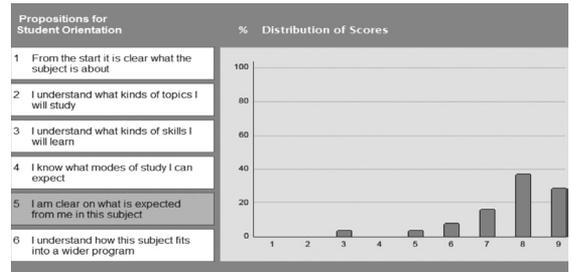
Figure 8: Learning Outcomes

Figure 9: Student Orientation

In Figure 7 for example, most of the group rated item 25 highly. Those who rated it lower had to consider whether they had taken the same steps as their fellow-students to read and absorb all the information made available to them.

Applications and adaptations

Early interventions

There is also scope for the lecturer to use this kind of process as an intervention, perhaps half way through a subject, to check how students are faring and to clarify matters or make minor adjustments. (In the masters programme I coordinate, organised as a series of three to four day residential schools, we run this process at the end of each residential.) In the example at Figure 8 from a group of 24 students, those who have given a lower score to item 33 may be struggling with the course material, or confused about the rules of the game, or unwilling to disagree with their more assertive classmates, or shy in group situations. Or with a group this size, there may simply not be enough airspace for everyone to have their say on a topic before the discussion moves on. In a setting where people are invited to declare how they experience a class, it is (sometimes) possible to sensitise more vocal students to the need of less vocal students for airspace, so they can find their voice when they have something to contribute.

Likewise, in the example at Figure 9, taken at an early stage of a subject, most but not all students were fairly clear about what was expected of them. Having heard what these students were uncertain about, it was not too difficult for the lecturers to give clearer guidance to these students on the matters they raised (in this case, mainly assessment tasks and assessment standards), so they could proceed without undue anxiety.

Larger groups

Meanwhile, for larger groups using a low-tech approach like the one outlined here, it would make sense to get a colleague or student to enter the data while the lecturer engages the class in a separate activity, and then play back and discuss results. (Alternatively, students can be asked to

email their results to the lecturer during the week prior to the class to allow for data entry before the class begins. But, the risk with this is that some students will forget to email their results, or do so but then fail to attend, or forget to bring their completed forms for reference, or forget why they scored things as they did a week earlier.)

Varying the survey content

Further, given control of the technology, the lecturer or programme coordinator also has scope to vary the content and wording of the survey items over time, in the light of experience. These things need not (and should not) be set in stone, since in a process designed primarily to identify localised problems and possibilities, statistical precision is not the main game. The main trade-off here is the one identified at the start of this paper, between a high level of context specificity (where flexibility to customise is a virtue), and a high capacity to benchmark student feedback results across contexts (where standardisation is a virtue).

While it makes sense to design this kind of interactive survey in a way that will mirror the main elements that standardised instruments already in use in one's institution seek to measure, it is at least as important to reflect the specifics of the resources available to people in this programme, the aims and requirements of this subject, the issues presented by this teaching and learning situation, and what this lecturer is trying to accomplish.

Conclusion

Given the institutional need for standardised and summative information for planning and reporting purposes, the kind of process outlined here can only supplement, rather than replace, the passive surveys that are already in use. Institutions need summative measures of their teaching and learning performance, not least because whenever they make changes to programmes, they are at risk of being accused of dropping standards or ignoring the needs of students or both, on the basis of impressionistic or anecdotal evidence.

But even so, there appear to be benefits all round in a supplementary approach such as this one. In the situation outlined at the start of this paper, where I received the results of the official student feedback survey undertaken some months earlier, there were no surprises or questions that I might have struggled to answer if my programme coordinator had asked me why I was getting this kind of result. Why? Because on that day in October, I also used the approach outlined here. When asked to complete the official quality of teaching survey, my class had already

engaged in a detailed discussion of this one. It is possible that doing this led to a better result in the official survey, since in the process outlined here, the lecturer is automatically seen as responsive to student feedback, and students have had a chance to share and discuss their views.

Finally, the process I have outlined here does not leave lecturers in a position where students can easily adopt the role of passive consumers of their education, or assume that the lecturer's primary purpose is to satisfy them by catering to every preference they express. As students, their relations with their lecturers, each other, and the wider institution are more complicated than this. As I have argued elsewhere (Sharrock, 2000) university students play multiple roles when they attend an institution to study for a qualification. They may be identified as customers, clients, citizens or subjects, depending on who and what they are dealing with at any given moment. As educators, we mediate a rather complicated set of institutional responsibilities in our dealings with students, blending detachment and concern as we balance student interests with those of third parties (such as employers) beyond the student-teacher transaction. To balance the range of obligations this relationship entails, we need supportive systems and discretionary space to define quality in ways that make good sense, and tools that are fit for purpose in a wide range of contexts.

Geoff Sharrock is Programme Director at the LH Martin Institute for Tertiary Education Leadership and Management, University of Melbourne.

References

Biggs, J. (2001). The reflective institution: Assuring and enhancing the quality of teaching and learning. *Higher Education 41*, 221-238.

Buchanan, J. (2011). Quality teaching: means for its enhancement? *Australian Universities' Review, 53*(1), 66-72.

Devlin, M. & Samarawickrema, G. (2010). The criteria of effective teaching in a changing higher education context. *Higher Education Research & Development, 29*(2), 111-124.

James, R. (2002). Students' Changing Expectations of Higher Education and the Consequences of Mismatches with the Reality. In Coadrake, P. (ed.), *Responding to Student Expectations*. OECD Paris, 71-83

Sharrock, G. (2000). Why Students Are Not (Just) Customers. *Journal of Higher Education Policy and Management, 22*(2), 149-164.

University of Melbourne. (2007). *Nine Principles Guiding Teaching and Learning.* Retrieved from: http://www.cshe.unimelb.edu.au/resources_teach/framework_teach/docs/9principles.pdf

UNSW (University of New South Wales). (2004). *Guidelines on Learning that Inform Teaching at UNSW.* Retrieved from: http://www.guidelinesonlearning.unsw.edu.au/