Evaluating Teaching Innovation

By Richard James

The evaluation of teaching and learning, and particularly of teaching and learning innovation, has become a priority in universities. The idea in writing this booklet is to prepare a short and succinct guide to important ideas about evaluation for staff who are evaluating new approaches to teaching, perhaps for the first time. The booklet draws on interviews with staff in various disciplines in the University of Melbourne who have led teaching and learning innovations. It includes brief case studies of a few projects and how they have been evaluated.

Arguably, universities are in the midst of a genuine revolution in teaching. The emphasis in the past decade on the quality of higher education has focused attention on the needs of students, and of identifiable student subgroups, and on how these needs can be best met. New information technologies are giving rise to new possibilities and academics are exploiting the potential of multimedia and the Internet. As well, universities are increasingly experimenting with alternative teaching strategies—such as problem-based learning and collaborative learning—to expand the range of learning outcomes and increase the relevance of student learning to future careers.

But innovations such as these do not always lead to improvement in student learning. This is why carefully designed evaluation is important—to determine in a rigorous way the influences of new curricula on student learning.

For the purposes of this booklet, evaluation of teaching and learning is defined as ‘a systematic attempt to determine the effects of an educational initiative, innovation or experiment on student learning’. Using this definition, the success of a teaching innovation will be determined principally by evidence of changes or improvement of some kind in student learning.

This booklet is based on the assumption that educational evaluation is a professional responsibility for academic staff, arising from a commitment to understanding the effects of teaching on students and to enhancing student learning. Evaluation is also a requirement in funded projects. Not all teaching innovation requires special funding, for some of the most useful innovations are simple, cost-neutral changes in style, structure and pace. But
wherever funding agencies are involved, such as the national Committee for University Teaching and Staff Development (CUTSD), which annually offers National Teaching Development Grants to initiatives aimed at improving the quality of teaching, or the University’s TALMET grants scheme, you can expect that evaluation is essential and that your proposed evaluation methodology will be one criteria on which projects are assessed. Funding agencies expect evidence that their money is being well-spent, and poorly designed and badly implemented approaches to evaluation may limit the chances of securing financial support or retaining future funding.

**Ideas about evaluation**

- In most educational settings, the purposes of evaluation are two-fold: to help steer projects towards a successful outcome and to assess and report the final project outcomes at an appropriate ‘end point’. The best evaluation designs will therefore include formative elements (those which feed information into project design) and summative elements (those intended to judge the final outcomes). Usually the information from formative evaluation is helpful for summative evaluation as well.

- If evaluation involves seeking student responses (as it often does) and if what they say is acted on, then this sends a clear message to students that their opinions are valued. This can not only enhance their learning, but also their satisfaction and involvement in the course.

- Despite the importance of evaluation, many academics do not fully understand what it entails. As one staff member commented, ‘evaluation is often very amateurish’. Evaluation is surrounded by a fair amount of mystique and there is a feeling among some academics that it’s a job for the professionals. ‘People are reluctant to spend the time and the effort to become good at it; evaluation can be very intimidating and sometimes boring’, was the comment of one staff member. The belief was also expressed that some teachers don’t have an interest in evaluation; they don’t see it as part of their role. Perhaps as a consequence, many academics say ‘you need professional support.’ Academic staff can make good use of external and professional support, either by seeking advice from the CSHE or the Multimedia Education Unit (MEU) in the planning and implementation stages, or by having CSHE or MEU staff directly involved in the actual evaluation by, for example, interviewing students.

- A common reason for introducing a change of approach is lack of staff time and resources to give personalised attention to students. It is somewhat paradoxical then, that innovations don’t always save the time of staff who implement them. Planning,
implementing, monitoring, making adjustments and evaluating the innovation quite frequently means more work, initially at least. If all of these are done conscientiously, it is going to demand staff time, and, generally, from more than one person. However, there are other rewards, particularly the satisfaction of doing something better, having more satisfied students, seeing students grasp concepts more readily. One staff member commented: ‘It doesn’t save staff time, but we hope we are teaching the same material better, in a different way. It gives students a chance to work interactively and in their own time’. Another found the extra administrative load of evaluating and keeping track of student concerns quite onerous at first, but commented: ‘You learn as you go along; I’m sure that with experience this administration would become easier’.

- Evaluation is sometimes not done, or not done thoroughly, because it’s regarded as a costly undertaking—and it can be. And because staff don’t know how to do it or don’t see it as important, it’s frequently not taken account of in initial costing. ‘It’s the unfunded bit in terms of time and money’, commented one lecturer.

- Be careful not to equate student satisfaction with better learning outcomes, although improved satisfaction is usually a positive outcome. Right from the early stages, be precise about what you’re evaluation is looking for. For example, what are the concepts in which you expect students will have improved understanding? Are you looking for the ability to transfer prior learning, or to apply concepts to the solution of practical problems? Are you looking for improvement in the skills of information search and retrieval? Evaluation should be closely linked to expected outcomes.

Equally, be alert to the unexpected. Many staff comment that evaluation ‘opens our eyes to a lot of things.’ It is sometimes difficult to remember that what appears straightforward and simple to lecturers may not be so for students: ‘We thought we had it right. Most of the shaping was done before the students saw it. The computer programmer, the instructional designer and the lecturers commented before the students used it. However, unanticipated errors and ambiguities were picked up by the students.’

- It’s a good idea to use more than one method to evaluate whether or not you have achieved what your objectives. But is it possible to over evaluate? As long as each evaluation strategy provides some useful information and you have the time, this is probably a matter of finding the right balance.

- It is important to share the results of innovative approaches. Don’t assume that what you have done is not worth communicating. Publishing helps others avoid having to re-invent the wheel.
Steps to effective evaluation

1. Timing

Most textbooks on evaluation emphasise the importance of building in evaluation procedures in the planning stages, or at least thinking about how you are going to evaluate early on in the project. In general, this is because it makes evaluation so much easier if you are aware of the need to collect, from the beginning, the type of information which you will need in order to be able to say whether the program ‘worked’ or not. Thinking about this in the early stages can also be helpful in clarifying what you actually hope to achieve.

On the other hand, it’s often important to do some actual evaluation, to get some feedback, early on in the project. The reasons for this are fairly obvious. As one lecturer said: ‘Evaluating throughout is important; it allows adjustments and improvements to be made.’ In computer based learning packages particularly, it is important to clarify things which students do not understand or find confusing, to rethink areas in which students are not responding.

At the same, it needs to be remembered that some of the processes involved in learning new material are very complex. Sometimes, even at the end of a year long course, it is ‘too early to tell whether students are achieving better’. Assessing better learning outcomes takes time.

It’s never too late to start. Not building in ways of collecting useful information from the beginning of the project may mean that there are some aspects you won’t be able to evaluate, however, it doesn’t mean that all is lost. There are generally some records, for example, of examination or test results which can be analysed or used as comparison measures, and some chance of getting retrospective views about the early stages of the innovation.

2. Methods

There are no hard and fast rules about the methods of evaluating new teaching and learning initiatives but there are, of course, some general approaches which are useful in a range of circumstances. Here are some common methods of evaluating, each with some brief comments.

Questionnaires are frequently used in evaluation but all too often the sole or the main focus is on whether or not students like the new course, or the parts of it with which they are satisfied or dissatisfied. While this is useful information (you probably wouldn’t want
to persevere with something students were very dissatisfied with, however good you thought it was), questionnaires can also be used to explore more complex, and perhaps more relevant issues such as what students are learning, what aspects are most useful, what could be improved. Depending on the purpose, questionnaires can be used at almost any time, but they may be particularly useful at the beginning and end of an initiative.

Questionnaires do have limitations. They are generally relatively structured and therefore place restrictions on the ways in which people can respond to them. Questionnaires usually reflect the compiler’s concerns rather than those of the respondents. One of our interviewees reflected on what he saw as another limitation: ‘Questionnaires have inherent limitations because they leave out the important dimension of body language. A questionnaire can’t tell me that some people use their time badly and sit there taking notes (whatever the method of learning they are confronted with).’

Straightforward comparisons are also often useful. These can be of:

- students’ exam or test results before and after the introduction of a new approach;
- standard tests of understanding of concepts in the area before and after the initiative is introduced;
- exam or test results of those who take part in the innovation and those who don’t—see ‘control groups’ below.

Following one of the innovations reported below, comparisons were made between assessment results from the new course and other courses run by the department, and the new course and other (second year) subjects. A somewhat more sophisticated comparison than test or examination results alone is a comparison analysis, conducted by experts, of the cognitive demand of essays and examinations in a new course and a previous course.

Sometimes as a means of comparison a formal experiment can be established with a control group. In instances where students volunteer to take part in a new initiative or a pilot program, those who don’t volunteer can be a built-in ‘control group’, a group against which to assess the performance or the responses of those who take part in the new program or approach. However, there are a few catches. It may be that there are prior differences between the volunteer and the non-volunteer group which are going to affect their performance. For example, it may be that students who volunteer are, on the whole, either the most or the least able, or the more creative, or the more willing to take risks. So it’s a good idea to check out these sorts of factors, and take them into account when analysing your comparisons. A lecturer who used a comparison of a volunteer and non-volunteer group noted that her own enthusiasm for the innovation may have influenced students to choose it and may have influenced student performance in the ‘experimental’ group.
Using control groups in education presents other difficulties as well. Usually it is impossible to ‘quarantine’ the two student groups, and some interchange of teaching materials can be expected. As well, students are likely to be critical, rightly, of any situation in which there is the possibility of inequity of opportunity. Consequently any controlled experiment needs to incorporate mechanisms to ensure that students are not disadvantaged in any way, particularly in grading outcomes.

Asking students to keep a journal of their experiences and their comments about the initiative can often be useful. Journals can be analysed for changes during the period, for example, increases in confidence and understanding. They can also reveal possible changes which need to be made along the way. In some circumstances, journals can be included as part of student assessment.

Student interviews are often a valuable means of evaluation, so long as the focus is clear. As interviews are generally time consuming, interviews of a small sample of students may be used in conjunction with other methods of assessing the response of the whole group. Interviews can be done with focus groups of students and these meetings often can provide opportunities to explore group responses in some detail. They can be arranged at appropriate times during the course. Of course interviews can be conducted by outsiders if you feel that impartiality or a critical view from the outside is necessary.

Informal observation of behaviour is useful in some circumstances, as long as your hunches about what it ‘means’ are followed up in some systematic way. Don’t forget that much formative evaluation also occurs simply through informal discussion with students and with staff involved in the program.

Straight mapping or counting of the number of students who turn up to voluntary sessions can provide some indication of whether students find them useful, although again, this needs to be followed up in more detail. Who comes along? Why? Who doesn’t turn up? Why not? What do the sessions provide or not provide?

Quite sophisticated data on student patterns of usage can be collected during computer-based instruction. Software can chart the time spent on various tasks, the areas where help is sought, and so on. An analysis of these files can often provide pointers to further questions about what works and what doesn’t, and most importantly, why it works or not, questions which can be followed up through formal interviews or informal discussions with students. Feedback sheets or opportunities to comment can also be built into computer assisted learning units.
One department, for instance, has compared the amount of time spent by students of different ability (as indicated by test and examination results) and found that students of similar ability spent varying time on computer assisted tutorials, prompting the observation that students had individual working styles and patterns, an important point to remember when planning teaching approaches.

So far, the focus has been on student responses and student learning, but the reactions of participating staff can also be a very helpful part of any evaluation strategy. In fact, involving other staff is generally advisable. Some academics we interviewed stressed the general importance of talking with, discussing and seeking the responses of academic staff who may not be directly with the innovation. The goals, possibilities and limitations of the project need to be communicated to other staff so that there is continued feedback, new ideas and reflection.

**Approaches used by University of Melbourne staff**

A number of the innovations described below involve the use of information technology. This reflects the large proportion of applications for teaching development grants which use computer-based and computer-assisted learning. The reasons for this are varied, reflecting the much greater acceptance of the technology generally, as well as its applicability in situations where universities are having to respond to a larger and much more diverse undergraduate population.

Peter McTigue and his colleagues developed computer-based learning (CBL) materials for first year chemistry students. Large first year classes, poor attendance at voluntary tutorials and economic imperatives, which made it difficult to maintain an adequate tutorial system, led to their introduction. Practical sessions were cut and replaced by workshop tutorials incorporating CBL. Evaluation included detailed interviews of students by a member of staff from the Centre for the Study of Higher Education; student questionnaires; reviewing student data files (done at the end of the course when staff had the time); general impressions from students and staff; and a retreat to review and plan for the following year.

Rick Willis was responsible for designing and implementing new feedback sheets for students of first year Biology. Large student numbers, the impersonal nature of the course, lack of staff resources, and lack of adequate feedback to students, were the motivating forces. The new feedback sheets provided more information, and made marking more efficient as far as staff were concerned. Evaluation included asking students and staff their opinions of the new sheets; asking students to rate the feedback sheets; looking at feedback.
centre for the study of higher education

sheets over two years; and comparing exam results after the introduction of the new feedback sheet with results from previous years.

Kevin Whithear extensively reorganised part of the second year course for veterinary students. Students now receive course information as interactive, multimedia databases, consisting of text and an associated library of catalogued digital images, movies and sounds. The motivation was largely pedagogical—the rapid expansion of knowledge, information overload for students, and concern to inculcate higher order skills rather than mere rote learning. A comprehensive evaluation included feedback on the design of the new approach from three representative groups of students; computer polls built into the software; questionnaires at the beginning and conclusion of the course; informal discussion with students; formal interviews with students; a cognitive demand analysis of essays and examinations in the previous and the new course; and correlation of assessment results in the new course with results in other parts of microbiology and other second year subjects.

Kevin Forward and his colleagues developed tutorials in engineering which take advantage of the Windows environment, and methods which allow linking of use of the tutorials and results into the management of student records. Motivating factors were a shortage of staff and a desire to provide something for students more akin to individual tuition. Staff and student feedback on a trial tutorial early in the development had provided useful information. Of particular interest was the unexpected and unintended learning from the experience; the process had taken the staff in directions they had not considered at the beginning of the project.

Graham Parslow was instrumental in developing computer tutorial support for students of Introductory Biochemistry. The motivation was a lack of staff resources for individual tutorial assistance. At the time of interview, twenty tutorials had been developed. They are ‘low level’ CAL (computer assisted learning) consisting of graphics and text but not multimedia. Evaluation included student feedback through a questionnaire; examination of data on student usage which is built into the software; informal observations of ‘body language’ in the lab; and monitoring of examination results over the period of use and comparison with results from previous years.

In Carol Johnston’s innovation, Dip Ed./B.Ed students (trainee teachers) volunteered to teach second year economics students in small groups. The motivation was to improve the depth of understanding of second year students (together with a desire for smaller group learning situations but lack of staff resources to provide it) and the perceived need to enhance trainee teachers’ ability to conduct small groups. Evaluation was comprehensive and quite thorough—a standard test of understanding of applied economics at the beginning and end of the year; questionnaires at the beginning and end of the year which focused on why students did or did not chose to take part in the micro-teaching/learning; assessment of
a journal kept by the trainee teachers; examination results; and periodic discussions and feedback with the group which allowed adjustments to be made as the course proceeded.