

# Evaluating the New Technologies

*[Ann Sefton](#)*

*Faculty of Medicine and Faculty of Dentistry, The University of Sydney*

## Summary

There is an accelerating trend towards the use of new technologies in teaching; the challenge is to demonstrate its effectiveness.

In approaching the evaluation of new methods, what were the aims of the initiators? While their over-riding expectation or hope is usually to enhance learning in a specific area, other possible outcomes may include reaching more students, providing experiences otherwise impossible, offering flexible access to a wider range of information, encouraging rehearsal and practice in virtual environments, the development of more generic skills - including the use of computers themselves. Unless these expectations are made explicit, the impact or effectiveness of the technological solution cannot be measured against its own goals. Ongoing evaluation in use can subsequently feed into quality improvement cycles.

In other situations, comparative judgements are sought but in many ways this approach is difficult. There is by no means agreement on the best methods of evaluation, even for the most basic of questions: Is the technology more effective in enhancing students' learning than are the alternatives it replaces? Is it cost-effective? Is it received better by the students? Conventional teaching methods have by no means always been evaluated rigorously, so the baselines for comparison are lacking or flawed. Complex variables (characteristics of the program in which the technology is embedded, students, teachers) inevitably confound any differences found, so absolute judgements are rarely possible.

Nevertheless, some strategies for evaluation have evolved and are often very effective within a local context. Examples from fixed media and web-based technology will be discussed.

## Introduction

New technologies are rapidly entering teaching and learning, yet in many cases the evidence for their effectiveness is yet to come. Evaluation of the technologies in enhancing learning is not easy, quick or cheap, yet we must do it well in order to plan appropriately. In thinking about evaluation, we have first to review what the educational purpose and rationale is.

The sorts of decisions that need to be made have different implications, depending on the size of the organisational unit that is concerned with the introduction of new technologies. At the broadest level, in the face of the technological revolution, institutions are increasingly forced to consider what their roles are: to remain campus-based with an emphasis on face to face experiences; to extend their reach using technology and rely on distance learning; or to embrace a mix of methods depending on local needs. The same institution might well distinguish between: undergraduate teaching; continuing professional education; postgraduate coursework

qualifications; and research degrees. Different strategies and solutions may be appropriate to meet the distinct needs of each of these groups.

Below the level of the institution, faculties are now developing policies on the uses of technology across their various degree programs. In other situations, decisions are taken by departments or by those organising individual units of study. Further, some freedom of choice may be available to individual members of staff making only a small contribution to part of a unit of study. The costs and consequences of each of those decisions are necessarily very different. A poor decision at the level of an institution or faculty will be extraordinarily expensive in time, money and in lost opportunities; a poor decision at the level of an individual member of staff may cost that individual and the affected students some angst. Being so close to the action, however, problems can be identified early and remedial steps can be more quickly applied.

Whatever the level of decision-making, the technical context is clearly critical for both staff and students. It may be easy to offer high quality access on a fast intranet on campus, but for evaluating resources that will be accessed at home, we must put ourselves into the relevant environments. Issues include the sort of technology in use, with the likely constraints on speed and accessibility, as well as the resolution of the screens if image quality is important. To that end, we need to know something about the situation of both the students and staff, including their access to computers off campus.

### **Evaluation of local strategies**

Whatever the organisational level, from the institution to the individual enthusiast, evaluation is essential, but its effective implementation may pose rather different sorts of problems. There are also issues that are common to the effective application of strategies for evaluation across the levels.

### **Meeting goals**

Those who are engaged in the processes of developing whole educational programs that are based on or supported by information technology and those whose uses are confined to a single unit or 'lesson' are aware of the high costs for development and infrastructure. They have particular concerns to ensure that their students' learning is of high quality and effective. At both levels, teachers want to see the extent to which the technology helps the students to meet the goals of the degree program and/or the individual unit of study. In order to do this, the goals or outcomes - whether of a whole program or a smaller unit - must be explicit and achievable.

As an example of program goals, The University of Sydney Medical Program makes strong statements about critical reasoning for medical practice, effective skills in communication and clinical examination, a capacity to understand community concerns and population issues as well as to behave ethically and to be reflective practitioners. A set of statements about the values of the curriculum includes student-centred, independent learning, reflection and self-evaluation, cooperation in groups, evidence-based decision making, effective skills in clinical work and in information technology. Clearly, the technological support designed for use in the program must be consistent with those aims.

While evaluation against a defined standard is ideal, in practice it is often difficult, particularly for whole programs or yearlong courses. Unless the outcomes and goals are clear - and they are not always so - it is impossible to test whether they have been met. Countless variables inevitably intrude - for example students may access many forms of learning, so the impact of any one component cannot be teased out. Often the starting point is not identified or measured, so gains are hard to quantify. Many of the specific or generic skills may be acquired from other studies or outside of the university. The timing of evaluation is also an issue: many goals or outcome statements for degree programs refer to qualities or skills that will be possessed by graduates well down the track, or throughout a lifetime.

Small programs or packages often have more modest aims - e.g. understanding one or a limited number of concepts, solving relatively simple but relevant problems, applying knowledge to a new environment, acquiring a variety of skills including skills in information technology itself, stimulating interest. Some of these outcomes can be tested relatively easily, but others are more problematical. Even with a smaller package, the designer might be aiming for longer-term gains in understanding and the application or synthesis of knowledge into the future, but for logistic reasons, observation and measurement is often limited to the immediate end of the period of study.

### **Understanding learning processes**

Teachers may also aim to understand something about the learning processes of their students from the ways in which they work through a program or package. Observing students at work on a program is expensive, but yields very valuable information about the quality of the program and its usefulness to the users. Thus, there is often an overlap between the specific evaluation of a component of information technology and more fundamental educational research. Obviously such approaches can generate important observations that go beyond the simple brief of evaluating a particular program or intervention. Such work is only in its infancy but in the longer term, the approach has the benefit of enhancing our understanding of how students learn differently - if they do - using technology rather than other tools.

### **Comparative studies**

Comparative evaluation sounds attractive, but is a very difficult task. It is easy to pose the obvious question: is this use of information technology more effective than alternative methods of teaching? It is much more difficult to find ethical and practical ways of answering the question unequivocally. Students learn in different ways, and value different sorts of experiences.

Some new approaches to 'evidence-based education' are, however, drawing on the ideas of the very effective 'evidence-based medicine' movement. Shortly, in several leading journals of medical education, an approach to the weighting of educational evidence will be reviewed and presented. I hope that this step leads to a vigorous debate and to a more secure understanding of the value of different sorts of evidence.

By analogy with the medical literature, randomised controlled trial designs that involve dividing students into two groups inevitably raise issues of fairness, the comparability between the groups

and indeed, the level of communication or collaboration between them. Further, it is difficult (and usually unethical) to contrive a placebo in educational terms. The common educational solution of comparing students in successive cohorts with and without an intervention introduces a range of uncontrolled variables. New strategies of meta-analysis have been applied in several educational cohort studies, however, suggesting one possible future direction. Unfortunately, an educational study cannot be carried out like an 'n of 1' trial of an individual with periods on and off a relevant drug (or program) since learning would be expected to be cumulative.

Any study design must access the students' responses directly. It is crucial to find out how students balance their learning between the range of educational activities undertaken; presenting students with a new strategy does not mean that they actually used it. Indeed, anecdotal examples give us pause for thought (e.g. some laser disks that were thought to have been spectacularly effective until it was realised that students found them so impenetrable that they resorted to small group self-directed learning in the library). Perhaps, in the end, we will come to rely on students' reports of their impressions of the comparative usefulness of a new tool for their own learning. We must aim to relate the evaluative information to students' learning styles so as to influence future educational design.

### **Quality improvement**

Another obvious use of evaluation is to feed back into a locally produced program for ongoing quality improvement. The cycle of evaluation can feed directly back into a process of re-design and enhancement. Such an approach requires that appropriate academic structures and mechanisms are in place to effect the agreed changes.

### **Evaluation of imported 'solutions'**

We may want to compare one technological intervention with another. On a small scale it may be readily apparent that one package is more effective or better received than another, but for larger scale decisions, the same caveats as were pointed out above apply.

With the burgeoning availability of information technology, our attention is increasingly drawn to disks or web sites that might provide useful teaching resources for our students without infringing copyright. We may be interested from the simple viewpoint of providing an illustration or a small teaching aid, or we may be considering the far more extensive adoption or adaptation of a whole unit of study or even an entire curriculum. Major developments worldwide suggest that programs from high quality institutions may become available for individual students or whole classes, some perhaps at affordable licensing fees. Given the acknowledged costs of developing even a fairly straightforward teaching unit, *not* re-inventing the wheel is obviously attractive. How, though, do we set about making judgements on the quality, likely effectiveness, ease of use and accessibility of learning packages or programs produced by others?

Once again, the key lies in identifying the aims of the program we are evaluating in the light of our own goals for the degree program and/or unit of study. What do we want the students to achieve at the end of the exposure to the 'lesson', unit or program? Is the design of the intervention compatible with both sets of aims? What relevant advances in learning will result

from the program, activity or aid? Is the level and style of learning appropriate to our goals? Is the educational design appropriate? What local modifications might we seek? Are they feasible and permitted by the licensing agreement?

Costs are always an issue: up-front charges and licensing agreements vary, making some highly desirable programs (on-line or on disk) unaffordable.

## **Designing evaluation of IT**

### **Teachers' concerns**

The power of the technology lies in the capacity to bring information together in engaging ways. Computers offer access to almost unlimited information. They can also offer simulations, replacing experiments impossible in the average classroom. For example, students can learn to patch-clamp individual cells in simulation, achieving with an inexpensive computer and program what would require hundreds of thousands of dollars worth of equipment. They can analyse their data on the same machines and present their conclusions using text, graphics and images. In biological and medical sciences, observations become possible that were once only available in well-equipped laboratories, either using software to mimic expensive recording and analytical machines or to provide access to simulations. With increasing ethical barriers to experiments on animals or humans, such replacements are proving invaluable. To be effective, though, such programs must challenge and interest the students who must be clear on the goals they are to achieve. The limitations are obvious - computers don't bleed - but well done, they can open access to observations that mimic first-hand experiences. They can provide plausible data for later analysis, presentation and discussion.

What teaching goal(s) is the IT intervention designed to meet? Are these aims appropriate for the overall program or unit of study? Goals/aims such as:

- learn a new concept;
- apply knowledge already learned;
- offer access to information and databases;
- solve problems;
- provide feedback on learning;
- rehearse a skill; or
- access/simulate/replace an experiment.

Before asking the students, some questions need to be satisfactorily resolved. Are the program's strategies and approaches consistent with our overall educational objectives? Do they:

- encourage active learning?
- stimulate problem-solving?
- trigger 'what if' speculation?
- stimulate student-student discussion?
- support further exploration?

- offer quizzes and/or feedback?

As teachers, we need to be assured that the program is built around essential and appropriate content or skills, that the level of difficulty for the particular student group is reasonable and that the intervention is designed to meet both specific and generic skills consistent with those of the program.

- Is the material relevant/important?
- Are the specific skills essential?
- Is the level of knowledge/skill right?
- Does it enhance useful generic skills?
- Is it well-matched to assessments?
- Are the outcomes consistent with program goals?

Other concerns lie with the technical issues and to some extent 'taste'. The importance of good educational design is only recently becoming recognised in tertiary education although its usefulness has previously been demonstrated in schools and technical colleges. The use of the information technologies has made issues of design and structure more apparent as we realise the need to match them with the overall goals and delivery of the programs. While some technical aspects can readily be analysed, issues of taste are more problematical. The generation gap often intrudes. That is, to me, a screen may be unreasonably cluttered, and the colours garish; to my students, iconic complexity is appreciated, indeed easily absorbed, and the colours are regarded as entirely acceptable or even 'trendy'. Thus issues of design and taste must be tested with the target population. Some simple observations, however, include:

- Is the program well constructed?
- Are the screens clear/acceptable to users?
- Is navigation easy?
- Are the instructions clear?
- Can the user exit easily?
- Is it too slow/fast?
- Does it include feedback?

Obviously, a program must be of high technical quality, without obvious quirks or glitches. It must be highly relevant to the group concerned to be maximally useful. These represent value judgements that can only be made in the context of the objectives of the overall program and of the technological intervention itself. Examples of fairly clear-cut issues of quality and relevance might be:

- Is the issue/task/learning important?
- Is the information accurate?
- Is the approach up-to-date?
- Are the examples appropriate?
- Are illustrations clear and relevant?
- If a simulation, how 'real' is it?

Central to the evaluation process is the experience and the views of the students concerned. Any process of evaluation must involve them directly, by providing adequate ways of tapping into their reactions and comments. First, though, it is necessary to establish some baselines.

- What are their expectations?
- What are their specific learning needs?
- What are their generic learning needs?
- Is the IT interactive and time-effective?
- Is the IT consistent with assessments?
- Does it offer helpful feedback?

Well designed questionnaires (on paper or on-line) and/or focus groups to follow up details, make it possible to gain access to their opinions. Other aspects are open to more objective observation - e.g. the number of times students choose to refer to a particular program out of class hours, the educational 'buzz' in the classroom when the program is in use, enhanced performance on standard tests. What do students and staff report?

Student views:

exciting, interesting? useful? boring? useless?  
easy to use? impenetrable?  
what aspects are valued for learning?

Objective measures:

does it stimulate discussion?  
do they seek it out frequently?  
can improved learning be measured?

For all of us in higher education, issues of costs become quite crucial. We need to be very clear in determining whether an intervention not only enhances learning, is well received by the students and staff, but is also cost-effective. Could the same result be achieved more cheaply? Discussions with those involved in the development of specific programs often suggests that the full costs are seldom calculated and made explicit. Educational and technological enthusiasts will cheerfully give up excessive time to development, over months or years, seldom counting the total hours and the costs of lost opportunities and other aspects of teaching neglected. The 'not invented here' syndrome leads to massive duplication as individuals are reluctant to accept a program from a colleague in another institution. In this important area, we must devise more effective means of communication to avoid such waste. Shareware programs, web sites and individual teaching items from databases like the National Teaching and Learning Database Project (including high quality images and diagrams) are becoming available, and we must learn to utilise these to our best advantage.

- Is the program effective and cost-effective?
- Is the students' learning increased?
- Do they enjoy the new learning?

- Do they rate it as high quality?
- Are they motivated to learn more?
- Are there cheaper but effective alternatives?

## **Using computers for program evaluation**

Networked computers can be used effectively to present evaluation questions to students at the end of a period of learning, after particular interventions or experiences, or when difficulties are perceived to be developing. Feedback to staff is fast and can be interactive. Numerical responses can be automatically collated, stored, analysed and reported in tabular or graphic formats.

As an example, in The University of Sydney Medical Program, every web page has a 'feedback button' so that comments can be fed to the relevant staff member(s) and to curriculum developers. Students study one clinical problem per week in the first two years: at the end of each week the problem-based learning groups feed back their overall impressions of the effectiveness of the learning experiences during that week. This feedback is expanded in intermittent focus groups attended by group representatives. Students have access to on-line self-assessment questions: they can, and do, email the relevant writer if they have difficulties. Throughout the program, students fill in on-line questionnaires on a wide range of different aspects. Such rapid and effective feedback is of great usefulness to staff in refining the curriculum, and is valued by the students. Another feature of the program is the 'staff site' which allows teachers to review all on-line teaching materials, outlines and timetables. This site offers a powerful opportunity for evaluation of the quality and relevance of materials presented, as well as providing information for teachers on the program itself.

## **Teaching students the skills of evaluation**

Encouraging students to approach evaluation seriously is important for their future educational development. Helping them to contribute usefully to the processes of evaluation often provides insights for them into the underlying structures and expectations of programs when these have not necessarily been made explicit. Good evaluation can turn students into enthusiastic collaborators with staff, enhancing collegiality and enriching the experiences of both groups.

More importantly, though, students will have to function in a future brimming with overloaded sources of information. The skills of critical appraisal and evaluation must apply not only to traditional texts or experimental data, but also to technology-based information. Information literacy is now being introduced in many contexts: the idea incorporates a complex set of generic skills - the capacity to locate, retrieve, use, modify, link, categorise, store and access information - applied to specific subject areas. It is increasingly essential for students to develop the ability to evaluate the quality, accuracy and usefulness of the computer based information they encounter. In the health sciences and increasingly in education, those interested in evidence-based practice are developing rational evaluative tools to judge the quality of evidence and perhaps similar approaches will develop in other fields. A conjunction is needed of those expert in information literacy (often librarians) and subject experts to contribute to the development of sophisticated critical skills in students. Only then will we meet the current and future needs of students who will be increasingly awash in a sea of information of very variable quality and applicability.



## **Conclusions**

Effective computer based learning:

- meets a clear educational need;
- is consistent with program goals;
- is cost-effective and timely;
- avoids errors, misconceptions;
- is set at the appropriate level;
- interests and motivates students;
- engages students actively in learning;
- is well-designed, easy to use; and
- encourages collaboration.

Good evaluation:

- is essential to meeting educational needs;
- must justify costs (time, money);
- is difficult, expensive, many variables;
- depends on explicit program goals; and
- is time-consuming and may not yield conclusive results.