Peter Ling
Learning and Teaching Services
Swinburne University of Technology
PO Box 218, Hawthorn, 3122, Australia
Phone 03 9214 5712
Fax 03 9214 8637
Email pling@swin.edu.au

FRAMES OF REFERENCE FOR EVALUATING EDUCATIONAL TECHNOLOGIES

ABSTRACT

This paper addresses issues in attempting to establish criteria against which to evaluate technological innovations in education. New educational technologies provide opportunities for gains in resource efficiency and in educational effectiveness. There are a number of available approaches to and forms of evaluation that might be applied to technological innovations in education – for example design evaluation, process evaluation, product evaluation and impact evaluation. This paper is concerned not with the phase or form of evaluation but with the derivation of criteria against which technological innovations in education may be judged at various phases of their development.

Four possible frames of reference are considered. One is oriented to the stated objectives of the particular educational innovation; one to comparison with alternative educational approaches; one to the benefits and costs anticipated from a knowledge of the state of the art of learning technologies; and one to criteria developed from a particular educational and/or social theory. These four frames of reference are interactive rather than exclusive. Each has a place in evaluation of educational innovations employing learning technologies. To engage all four, exposes contests between the theoretical foundations of an innovation, public policy, local pragmatics and individual objectives. Evaluation in this context, then, identifies dialectical relationships and provides a transformative tool for the construction and reconstruction of technology in education.

The paper reports an attempt to employ the frames of reference to evaluating an innovation that combined the introduction of problem-based learning with Web-based tuition in a course conducted at an Australian university

FRAMES OF REFERENCE FOR EVALUATING EDUCATIONAL TECHNOLOGIES

Peter Ling
Swinburne University of Technology

In a world of technological development evaluating innovations employing educational technology is a challenge confronting both educators and educational administrators. New educational technologies and new uses of existing ones provide opportunities for gains in resource efficiency and in educational effectiveness. Are the anticipated benefits realised and are the benefits worth the effort? What of the wider implications such as the social impact of technologies in a knowledge-based society? This paper focuses not upon evaluation instruments or methodologies relative to these matters but upon how to establish the criteria for evaluating innovations employing new learning technologies. There are several available frames of reference for evaluating educational technology.

SOME ISSUES

Evaluating innovations employing educational technology raises a number of issues and challenges, some of which relate to evaluation of educational technology in general and some of which derive from focusing on innovations.

With regard to the latter, the boundaries of the types of innovation to be evaluated need to be defined. Use of educational technologies can be more or less innovative or innovative in some contexts while established in others.

The other challenge is to define educational technology. In trying to evaluate the impact of innovation using educational technology do we include new uses of old technology; do we include low-tech innovations, such as new print materials, as well as high-tech? The use of educational technology does not imply any particular learning process informed by a particular learning theory. It could involve teacher driven construction, based on behaviourist theory or on information processing precepts. On the other hand it may be premised on individualisation of learning consistent with humanistic or with cognitive structural approaches. It may involve design catering for graded or alternative conceptualisations. The student may be largely passive or active and learning may take place through discovery or through direct instruction. Educational technology is not susceptible to an embracive evaluation. In fact the point is to evaluate a teaching and learning innovation which involves educational technology, not to evaluate educational technology, *per se*.

There are alternative approaches to establishing criteria by which to evaluate innovations employing educational technology.

CRITERIA

Goals or Objectives Based Criteria

If the goals or objectives of the project can be identified they are frequently used to provide the criteria for evaluation. This approach is a satisfactory approach if the objectives of the particular innovation have been accepted as being worthwhile, if any unanticipated outcomes are seen as irrelevant and if costs are seen as given and acceptable. This is, at least *prima facie*, the case for some specially funded projects, such as projects funded by a university or by an outside agency which calls for submissions for projects meeting certain criteria and provides support to a predetermined level. Even in these situations, evaluating a project against its own objectives is limiting. This method of evaluation will give no indication of unanticipated outcomes that may be as educationally or practically important as the intended outcomes. It does not allow for shifts in objectives. Focusing on project objectives can lead to costs being ignored or at least taken as given, being those specified in a project submission. In fact there are likely to be costs, both direct and indirect, which were not identified in project submissions. To ignore extraneous benefits, costs and other effects limits understandings that could inform future actions.

Comparative Criteria

An alternative approach is to compare the outcomes of an initiative with the pre-existing condition of the learner or with the effectiveness of an existing approach to the learning task; that is to compare learning under the innovation to a prior or concurrent approach that does not employ the media innovation.

There are two sets of issues in attempting a comparison based on the latter approach. The first is how it *could be* operationalised? What needs to be kept constant for comparative purposes? If the innovation had very wide application it could be trailed in a variety of circumstances and an inferential approach taken to its outcomes in comparison to traditional approaches, provided a consistent measure of outcomes was used. Wide application is not generally available which suggests the alternative approach of trying to keep as many factors constant as possible. This is likely to prove difficult. Cohorts of students will not be identical. Innovations are complex, changing many elements at once. The learning environment is likely to change in multiple ways.

The second set of issues in comparing old approaches with new educational media approaches is whether there is a commonality that *should be* compared. It would be unusual for educational and training initiatives using new learning technologies to simply change the form of teaching without impact on other aspects of the educational transaction. Initiatives are likely to arise from reconceptualisation of learning in the subject area as much from the availability of new teaching technologies. The nature of teaching and the role of the teacher may change from instructing expert to guide, facilitator, mentor, fellow learner, resource manager, while the role of the learner may change from passive recipient to client, explorer, problem solver, creative manipulator, cognitive apprentice, or evaluator. Even what constitutes the subject area, the definition of its boundaries, relationships between disciplines, who owns the knowledge, who can add to it, who can challenge it, all become open.

It would not then be appropriate to measure the success of approaches using new learning technologies against traditional approaches by using traditional student assessment tools such as a standardised test based on a set text. One could expect the objectives, processes and learning experiences to differ. Take for example anew learning technologies innovation which replaces expository teaching derived from a notion of transfer of information with discovery-based learning which values development of discovery techniques as much or more than the information acquired in the process. To apply a test of information acquisition alone as a means of comparison between approaches would be inappropriate.

Criteria Based on Conventional Wisdom

A further alternative in establishing criteria is to use benefits and costs anticipated from the literature or experience as a basis for evaluation. The result is to evaluate innovations employing digital technologies against benefits which could, on the basis of a conventional wisdom, be expected to flow from them and to likewise evaluate them on the basis of costs which could be expected to be incurred.

The rationale for this approach is that it can take into account a wider range of benefits and costs than those identified by the designers of a particular project. It takes on board broader experience in the field. It also gives a basis for a broader comparative evaluation than that available when individual project objectives are used to evaluate new learning technologies innovations. Expectations that arise from conventional wisdom may be that innovations will: provide new educational experiences; offer greater options for student selection of learning activities; extend information resources; extend opportunities for exchanges between students and between students and teachers; provide better opportunities for monitoring individual student progress; and provide wider access to learning.

On the cost side, costs relating to hardware requirements and staff time required for development could be anticipated along with some training and facilities costs.

Theoretically-Based Criteria

This approach to evaluation criteria requires, not just a theoretical position on evaluation, but a theoretical position in relation to the realm being evaluated. For example where learning outcomes are evaluated you need a theoretical understanding of learning processes. If organisational costs and benefits are an issue then a theoretical position on management is required. Where access to education is an issue a position on equity is needed.

To take learning theories as an example, the criteria that one would use to evaluate an education or training innovation would differ according to the understanding of learning held by the evaluator. If you have a behaviourist understanding of learning you might look for a systematic, step by step approach, with frequent testing that resulting in positive or negative reinforcement as appropriate to produce the prescribed learning outcome. An evaluator with a constructivist view of learning might be looking for the opportunity for the learner

to engage with the material, bring personal experiences and needs, apply their own meaning and emerge with something applicable to their own situation.

As for any of the approaches to establishing criteria the theoretical approach has its limitations. It may not address issues of the accountability of innovators to meet agreed objectives, in itself it says nothing about the comparative value of the innovation, and it will not be suitable for a decision maker who holds different ideals or theoretical understandings to those of the evaluator.

Frames of Reference for Selecting Evaluation Criteria

A model for selection of evaluation criteria based on the alternative frames of reference outlined above is tabulated in Figure 1. Each of the approaches to establishing evaluation criteria has its use and its limitations. The approaches can be combined. For instance for accountability purposes one may be obliged to take an objectives based approach to evaluating an innovation yet still wish to report on unintended outcomes. In combining criteria purposes need to be consistent or at least, if they are inconsistent, then the inconsistency needs to be highlighted when reporting.

A CASE STUDY

The issues outlined above were confronted in attempting to evaluate an educational technology initiative in computer science at the Royal Melbourne Institute of Technology (RMIT). Staff in the Computer Science Department observed that students undertaking a study of artificial intelligence (AI) using lecture and text book inputs and supported by tutorials had a number of difficulties. They had difficulty in gaining an overview of the subject or even of particular topics such as *search* or *logic*; they focused on the detail. Students failed to see the application of the topics to real life situations. The problems they dealt with in practical exercises seemed apart from the world of work for which their course was intended to prepare them. Teaching difficulties were compounded by the fact that the AI subject could be undertaken by students with a range of academic backgrounds in computer science.

The staff proposed an alternative approach to give students a framework for exploring the subject which would allow them: to locate their learning within the subject as a whole; to relate the subject to possibilities for application in the real world; and to have greater facility to move in the direction and at the pace preferred by the individual student or by groups. The two basic premises of the approach taken were to generate learning from the exploration of problems; and to allow, within some specified expectations, for individual or team decisions about the areas to be explored and the pace and depth of the exploration. Features of the approach adopted were the generation of learning from a focal problem; the location of the learning in a conceptual map of topics; and the use of computer based learning located on Internet to present problems, to provide learning resources and to track progress.

A focal AI problem - the operation of a fully automated taxi system - was used with a view to prompting students to envisage the types of problems faced in developing AI systems, that is to obtain their own overview of the subject. The idea was to allow students to then define and explore sub problems. Texts, references and course materials placed on Internet provided some learning resources, though students might also use their peers, experts and tutors as resources. The computer based learning program comprised two related elements: a course web of learning resources constructed by topic but accessible though hypertext and a problem web which assisted students - as they subdivided the focal problem into sub problems - to locate helpful learning resources housed in the course web.

Evaluation was used at each stage to observe student behaviour with elements of the program and to get feedback on attitudes, difficulties and preferences of students. For example, prior to the construction of any element of the computer program students were broken into small groups (of three) and asked to define sub-problems which they would need to address if confronted with the focal AI problem - the operation of a fully automated taxi system. Three levels of tutor intervention - tutor directed, tutor assisted, and tutorless - were employed with different groups as they worked with the problem. The groups varied in ethnic composition, gender, and computer science backgrounds as well as in terms of tutor intervention so that likely causes of differences between them were hard to isolate. Outcomes, however, in terms of the capacity to subdivide the focal problem outcomes were encouraging.

Plans for summative evaluation focused on a number of possibilities including evaluation against the project objectives and evaluation against the previous approach. It was intended that the evaluation should take an open format allowing for unexpected outcomes.

A number of observations can be made about the evaluation plans. Firstly that there was an undertaking to evaluate but no detailed plan for evaluation at the time the project submission was drafted. To some extent an evaluation plan has to emerge and adapt as projects move through stages form concept to submission to design to trialing, etc. Secondly the form of the project objectives did not allow them to readily translate into evaluation measures. Thirdly costs and disadvantages were not originally envisaged as a component of the evaluation. Fourthly having a control group to compare with was not envisaged, as it is not feasible to take the old and new approach simultaneously. Fifthly the staff initiative in establishing the project indicated a change in thinking about both how students might best learn the subject and what it was important to learn. In particular it placed more importance upon learning process and less upon learning outcomes in terms of information acquisition. It placed more importance upon appreciation of the big picture and less on detailed manipulations. It placed more importance on application to world-of-work problems and less on set exercises or on games type problems which can be a feature of AI studies. With these shifts it makes it inappropriate to use a single tool such as the traditional end of semester examination to test student learning under both the former approach and the innovative one.

Three of the four frames of reference tabulated above were employed in the evaluation of the project. The selection was driven by the imperatives to improve learning; to confirm project objective were met; and to satisfy stakeholders that the new approach was at least as successful as the previous one measured in traditional terms. The project was not evaluated against the educational potential of the medium.

The project objectives which were used to provide one set of evaluation criteria were to:

- Improve staff expertise in problem based learning
- Develop a set of contextualised problem based learning materials in AI
- Provide a method of building up course materials supporting contextualisation
- Improve learning outcomes

It was agreed that evaluation should at least in part attempt to compare the innovative approach with the established approach and that assessment of students would contribute to that evaluation. Due to concerns with security of individual student assessment and having regard for external expectations it was decided that assessment should include an examination on the traditional subject matter. As a corollary it was decided to incorporate some element of the problem solving approach taken in the project in comparing student performance between the two approaches.

For design evaluation purposes and for formative evaluation purposes criteria derived from learning theory were applied including: provide for interaction student/student and student/teacher; allow experimentation with concepts; allow collaborative learning; provide authentic challenging problems related to pertinent concepts; and allow for multiple entry conditions & multiple attainments.

Instruments employed to carry out this form of evaluation were staff observations of student behaviour, interviews with students and questionnaires to probe the influence of the innovation on learning.

CONCLUSIONS

Evaluation requires a frame of reference. Educational technology innovations often occur in a context of multiple obligations requiring more than one form of evaluation and implying more than one frame of reference. In the case study the innovation was funded on a project basis with an obligation to report against the objectives specified in the project submission. It replaced a traditional approach in which staff, students and agencies in the field had an interest. The stakeholders had to be satisfied that attainments that they valued in the old course were still being delivered. At the same time the incentive for the innovation was some new understandings among teachers of effective teaching/learning transactions in their discipline and success needed to also be reckoned in these terms.

Footnote

* The author wishes to acknowledge the work of the RMIT Computer Science AI Project Team:
Dr Lawerence Cavedon, Dept. of Computer Science
Dr Vic Cieselski, Dept. of Computer Science
Mr Daryl D'Souza, Dept. of Computer Science
Dr James Harland, Dept. of Computer Science
Mr Steve Hages, Dept. of Computer Science
Mrs Shelia Howell, Dept. of Computer Science
Mr Wayne Jenkins, Dept of Computer Science
Dr Lin Padgham, Dept. of Computer Science

References

Meyers, N.T. (1979). Gold or dust? information technology in higher education. *Australian Society of Educational Technology, Yearbook*, v4, pp45-54

Montgomery, A.Y. (1988). Information technology in higher education: new frontiers. *ASCILITE '88: Computers in learning in tertiary education*, Proceedings of the sixth annual ASCILITE conference

Pence, J.L. (1992). Transforming campus culture through resource based learning. New directions for higher education, v22, n2, pp113-122

Tilidetzke, (1992). A comparison of CAI and traditional instruction in a college algebra course. *Journal of computers in mathematics and science teaching*, v11, n1, pp.53-62

van der Molen, D.J.L. & Predebon, D.A. (1995). The design and application of computer programs for teaching of engineering theory and design', *ASCILITE '95: Learning with technology*, Proceedings of the thirteenth annual ASCILITE conference

Figure1 Frames of reference for evaluating educational innovations

Evaluation phase	FRAMES OF REFERENCE FOR EVALUATION		
	Evaluation imperative	**Evaluation basis**	**Evaluation criteria**
Needs analysis & Market analysis	**[1] Objectives based criteria**		
	Project accountability for meeting objectives	Measured against objectives	Depend upon objectives
Design evaluation	**[2] Comparative criteria**		
	Stakeholder expectations of standards being improved or maintained	Comparative (compared with the former condition of the learner or with a former approach)	Equal or better performance compared to previous condition or previous approach e.g., equal performance on traditional subject examination
Formative evaluation	**[3] Criteria based on expectations of technology**		
	Delivering expected benefits	Measured against conventional wisdom regarding knowledge media	Measured against expectations of benefit media can deliver & associated costs e.g., self paced learning, alternative entry & exit points, tracking of student progress, easily accessed and operated by the user
Summative evaluation	**[4] Theoretical criteria *e.g. A theory of learning***		
	Improving learning:- e.g. Allowing negotiation of goals & assessment e.g. Allowing student & teachers to articulate understandings & explore consequences e.g. Relating to the learner's world	Informed by learning theories	Provides for interaction of student/teacher Allows for multiple entry conditions & multiple attainments Allows student interaction with content to create concepts Allows experiment-ation with concepts Allows collaborative learning Provides authentic challenging problems related to pertinent concepts