

# **An introduction to the Evaluation of Learning Technology**

**Martin Oliver**

Higher Education Research and Development Unit  
University College London, 1-19 Torrington Place  
London, WC1E 6BT, England  
Tel: +44 20 7679 1905  
[martin.oliver@ucl.ac.uk](mailto:martin.oliver@ucl.ac.uk)

## **ABSTRACT**

Evaluation can be characterised as the process by which people make judgements about value and worth; however, in the context of learning technology, this judgement process is complex and often controversial. This article provides a context for analysing these complexities by summarising important debates from the wider evaluation community. These are then related to the context of learning technology, resulting in the identification of a range of specific issues. These include the paradigm debate, the move from expert-based to practitioner-based evaluation, attempts to provide tools to support practitioner-led evaluation, authenticity, the problem of defining and measuring costs, the role of checklists, the influence of the quality agenda on evaluation and the way in which the process of evaluation is itself affected by the use of learning technology. Finally, these issues are drawn together in order to produce an agenda for further research in this area.

## **Keywords**

Evaluation, Learning technology, Practice

## **Introduction**

Evaluation, put simply, is the process by which people make value judgements about things. In the context of learning technology, these judgements usually concern the educational value of innovations, or the pragmatics of introducing novel teaching techniques and resources. Less frequent, but still important, are judgements about the costs of such innovations. (Judgements about 'worth', as opposed to 'value', in the terminology of Guba & Lincoln, 1981.)

This article summarises current issues concerning the evaluation of learning technology, with a focus on Higher Education. As well as acting as a general introduction to the area, it will provide a context for the articles that form the basis of this special issue. These have been selected from the Evaluation of Learning Technology 1999 conference, which took place in London (Oliver, 1999a), and represent some of most important topics discussed at the event. (The full range of topics is presented in Oliver, 1999b.)

In order to provide an adequate context for these articles, this article will open with a summary of relevant trends in the broader evaluation community. This will be followed by a section in which these themes are related to learning technology, and important developments are illustrated. Finally, conclusions will be drawn that highlight which of these areas are likely to remain important in the future.

## **Issues facing the evaluation community**

The focus of most mainstream evaluation journals is social policy, often concentrating on the implementation of governmental initiatives. Although there are many concerns specific to the context of learning technology, as a newly emerging discipline, it is still grappling with problems that have been debated for decades within these mainstream fora. For this reason, it is useful to open with a discussion of this broader context.

Perhaps the most long running discussion within the evaluation community is the 'paradigm debate'. This focuses on the merits of qualitative versus quantitative evaluation methods. On the one hand, quantitative methods claim to be objective and to support generalisable conclusions. On the other, qualitative methods lay claim to flexibility, sensitivity and meaningful conclusions about specific problems. Quantitative evaluators challenged their colleagues on the ground of reliability, sample validity and subjectivity, whilst qualitative practitioners responded in kind with challenges concerning relevance, reductionism and the neglect of alternative world views.

Although this debate has continued for decades (see, e.g., Light & Smith, 1970), in recent years, discussion of these topics has started to wane. Several factors can be argued to have contributed to this development. The first, simply, is that neither side was able to put forward conclusive arguments. The implication of this is that there is no 'magic bullet' for evaluation (Oliver & Conole, 1998a). Instead, different methodologies will be useful depending on the situation in which the evaluation takes place.

This is closely related to a second important theme. Within the evaluation community, a new philosophy has emerged that eschews firm commitments to any one paradigm in favour of a focus on pragmatism. Rather than having a theoretical underpinning of its own, it involves a more post-modern view that acknowledges that different underpinnings exist, and adopts each when required by the context and audience.

This approach is well illustrated by Patton's utilization-focused evaluation (1997). Central to this view is the idea of evaluation as a means to an end, rather than an end in itself. Methodological concerns about validity, reliability and so on are considered secondary to whether or not the process helps people to do things. Patton provides various examples of real evaluations that have been perfectly executed, are well documented, but have sat unread on shelves once completed. In contrast, he also illustrates how "quick and dirty" informal methods have provided people with the information they need to take crucial decisions that affect the future of major social programmes.

Utilization-focused evaluation, like participative evaluation (Patton, 1997), also illustrates a third theme. This is the shift of power away from the evaluator as an arbiter acting for the commissioning authority and towards evaluation as a collaborative process of building mutual understanding. Perhaps the ultimate expression of this is action research (Schön, 1983), which sets out with the aim of empowering practitioners by providing a framework that enables them to carry out evaluations on their own. One important consequence of this shift is that it adds an educative element to the process of judging value.

*We believe that, to varying degrees, each of these evaluation approaches positions the evaluator as a facilitator of learning where stakeholders and program participants learn about themselves and each other, and the programme, through their involvement in the evaluation process... In this context the evaluator seeks to teach clients and stakeholders evaluation skills and processes so that they may continue to engage in evaluation practice when the 'evaluator' has left the scene.*

*(Preskill & Torres, 1999, p. 55)*

This philosophy underpins the notion of evaluative enquiry as a strategy for organisational learning (Preskill & Torres, 1999). This approach advocates a continuous process of evaluation by practitioners (as opposed to evaluators) as a strategy for dealing with organisational change and meeting the need for responsive, adaptive organisational structures. Interestingly, this philosophy also represents a shift from learning as a process of informing individuals (such as those in positions of power) to a socio-cultural model of learning (e.g. Brown et al., 1989).

*Learning from evaluative inquiry is a social construction that occurs through the involvement of multiple constituencies each of whom represents different perspectives. It is socially situated and is mediated through participants' previous knowledge and experiences.*

*(ibid, p. 44)*

Moreover, this style of evaluation represents a shift in focus away from self-contained programmes or projects to the ongoing evaluation of processes and systems. It requires the creation of a culture of reflective practice similar to that implied by action research, and has led to research into strategies for efficiently communicating and building knowledge (Torres, Preskill, & Piontek, 1996).

A final theme for this section is that the status of evaluation has been challenged by processes with similar aims. For example, the relative merits of evaluation and auditing have recently been discussed (Gunvaldsen & Karlsen, 1999), and the relationship between evaluation and quality assurance procedures has started to be investigated (Gilroy et al., 1999).

To take one example of this, performance management has recently been compared with evaluation research as a way of making programmes accountable (Blalock, 1999). In this discussion, the two approaches are contrasted in terms of performance management's emphasis on the outcomes of results and evaluation's focus on

understanding how events have come to pass. Given the forum in which this debate occurred, it is not surprising that much of the discussion focuses on the shortcomings of performance management, such as:

- Its inability to attribute cause to effects. (Few performance management systems can demonstrate that outcomes were not coincidental, or the result of some parallel activity.)
- Concerns about the validity of measurements. (Do the measures commonly adopted in performance management capture any important information about the processes being studied, or are convenient quantitative indicators being used in the place of complex, messy, qualitative information?)
- Issues of emphasis and oversimplification. (Performance management leads to a culture of “what gets measured gets done”; by implication, what does not get measured may be neglected – even if it is important.)

Shortcomings of evaluation research are also identified, however, such as the difficulty of making information available in a timely, ‘at need’ manner, and its relative expensiveness. The predictable outcome of such debates is the finer delineation of situations well suited to each approach, and the recognition that the different procedures may, in many ways, be complementary. In many ways, this appears to follow the same pattern as the earlier paradigm debate, and it seems likely that it will conclude in a similar way.

Whilst this brief synopsis of selected themes cannot hope to do justice to the breadth and depth of debate within the community, it does help to provide a context in which the evaluation of learning technology can be considered. Having provided this groundwork, this specific instance of evaluation will be considered in its own right.

## **Learning Technology**

Learning technology is an area with many names but few definitions. It can encompass educational multimedia, web-based learning, computer-assisted learning, and a whole host of other related topics, and is usually understood to be synonymous with Educational Technology (Oliver & Bradley, 1999). Irrespective of the terminology, all of these terms focus on the use of technology to support innovations in teaching and learning.

In the context of evaluation, this immediately provides a scope and an agenda. Evaluation becomes focused on issues such as the value of such innovations, the pragmatics of introducing them into the mainstream and the ethics of testing them out on real students. Moreover, a pragmatic emphasis arises from the fact that much learning technology research and development in Higher Education is funded in the form of self-contained projects. A common condition on such funding is that the project team should be able to demonstrate that they have achieved their aims, usually through some kind of evaluation (e.g. Tavistock, 1999).

Given this context, it is possible to review recent discussions and identify current issues facing the evaluation of learning technology. In this section, a selection of these is presented, including themes unique to learning technology as well as those that reflect the concerns of the wider evaluation community.

### **The paradigm debate revisited**

Just as mainstream evaluation has recognised that different methodologies have their own strengths and weaknesses, a similar position is now accepted within the context of learning technology. Several authors have advocated using qualitative and quantitative methodologies in order to triangulate results (e.g. Jones et al., 1996), thus enhancing the credibility of evaluation findings (Breen et al., 1998). Such models have been described as hybrid approaches (e.g. Oliver & Conole, 1998b). Other important factors that have contributed to this development include the adoption of utility-based approaches from mainstream evaluation (e.g. Tavistock, 1999) and the description of the different strengths and weaknesses of approaches (e.g. Harvey, 1998).

However, the paradigm debate is not entirely dead. An evaluation of an evaluation conference on learning technology identified the position of the two keynote speakers, one of which focused on qualitative methods and the other on quantitative, as being the single most divisive feature in terms of participants’ feedback (Oliver, 1999c). In part, this reflects the diverse backgrounds of researchers within learning technology. As an inherently multidisciplinary research area, it involves academics from a wide range of disciplines (such as education, psychology, computer science, etc.), each of which has its own traditions, values, criteria and practices (Becher, 1989). Consequently, it should come as no surprise that different members of the learning technology research community find some methods more persuasive than others. (An important consequence of this is the realisation

that one criteria for deciding on the usefulness of a methodology will be whether or not the audience will find such an approach persuasive, or even acceptable.) Although a consensus may have been reached amongst evaluators, sensitivity to this issue will continue to be important.

### **Evaluator or practitioner?**

The link between research funding and the need for evaluation has led to many practitioners carrying out their own studies. This is a marked contrast to traditional mainstream evaluation, which has traditionally involved an external expert being called in to carry out the study, although it does push the evaluation of learning technology towards more inclusive approaches.

Several issues have been identified that relate to this situation. It has been argued, for example, that bringing in an external expert on behalf of the commissioning group may allow a more professional, less biased study. However, evaluation by the team involved also has advantages, such as timeliness, understanding of the innovation and access to data (Bradbeer, 1999). Another important concern is that practitioners, who may be experts in a range of areas, may not have any prior experience of carrying out an evaluative study (Oliver & Conole, 1998a). In addition, evaluation often forms a relatively low priority within projects, not least because it is an unfamiliar and poorly supported activity that can seem unrelated to the completion of other, more tangible, project aims. Such perceptions can jeopardise the value and even the completion of the study (Oliver et al., 1999b).

The nature of funding in this area means that the issues outlined above cannot be solved; they can only be borne in mind. Many projects will have to conform to the evaluation requirements of the funding body, leaving little room for arguments over who might be best placed to carry out the evaluation in an ideal situation. One pragmatic strategy for coping with these problems focuses on the most appropriate type of evaluation that each type of evaluator can undertake. Bradbeer (1999), for example, suggests that internal evaluators should focus on formative evaluation, leaving the summative evaluation to people external to the programme. Another approach involves the development of resources that can provide support for practitioners, as discussed below.

### **Tools for practitioners**

The need for practitioners to carry out their own evaluations, as outlined above, has led to concerns about expertise. Lecturers, for example, may have expertise in their discipline and in teaching, but it is unreasonable to assume that they will have expertise, training, and in many cases even experience of carrying out programme evaluations (Oliver & Conole, 1998a). As a consequence, several tools have been developed to support practitioners engaging with evaluation. *The Evaluation Cookbook* (Harvey, 1998), for example, provides a series of evaluation 'recipes', each summarising a methodology in an easy to follow form, complete with hints and tips. These have been contributed by authors with expertise in using that particular approach, providing a rich and accessible knowledge base for practitioners to draw upon.

Whilst the cookbook provides a 'how to' guide for implementing evaluation studies, the ELT toolkit (Oliver, 1999d) focuses on their design. It is structured around a model of evaluation design that incorporates six stages, with the first two and the last relating to the context, and the middle three focusing on the details of the study itself. These steps are as follows (Oliver et al., 1998):

1. Identification of stakeholders
2. Selection and refinement of evaluation question(s), based on the stakeholder analysis
3. Selection of an evaluation methodology
4. Selection of data capture techniques
5. Selection of data analysis techniques
6. Choice of presentation format

At each step, users complete activities that support their use of a structured knowledge base (consisting of descriptions of methods). This is argued to help them make informed decisions that would otherwise be beyond their existing level of expertise. The toolkit is based on the pragmatic framework for methodology use outlined above, in that no single approach is presented as being correct. Instead, all are described in terms of their distinguishing characteristics, and the purpose of the exercise is to allow practitioners to select the approach best suited to their current situation.

An integrated approach that covers both design and application is the Flashlight programme. The Flashlight project has produced a questionnaire-based toolkit that provides a simple structure for evaluation by practitioners. The tool is based on an analysis of three elements:

- A technology
- An activity for which it is used
- The educational outcome of the activity

The tool also provides guidance on what it considers the characteristics of a good evaluation to be. These include studies of situations that are:

- Enabled or supported in crucial ways by the technology in question
- Repeated over time (in varying forms)
- Repeated in different courses (in varying forms)
- Exceptionally educationally important
- Exceptionally focused

Although these may be useful in determining whether a particular evaluation context is good or not, it focuses on choosing situations that suit the Flashlight methodology, rather than on selecting an approach that suits the situation.

The bulk of the tool is concerned with identifying questions that can be used to generate data. This includes consideration of five thematic prompts (e.g. “questions about the use of the technology”), and access to the Current Student Inventory, a repository of questions devised by other users. Based on a brainstorming exercise, questions are then created (possibly based on or incorporating examples from the Inventory) that can be used in surveys or structured interviews. Although interviews are mentioned, the focus for the tool is, quite clearly, questionnaire design.

The Flashlight tool is based on the premise that “very different educators need to ask similar questions” (Ehrmann, 1999). This position has been challenged as an over-simplification (Oliver & Conole, 1999), albeit a useful one. This assumption allows the Flashlight project to justify a focus on only one type of methodology, but it ignores the variety of more appropriate methodologies that the ELT toolkit, for example, seeks to identify (Oliver & Conole, 1998).

Although criticisms of the project can clearly be made, it is important to recognise that the tools it has provided a focused on usefulness and usability, and have been employed successfully in a range of real settings (Ehrmann, 1999).

A comparable approach involves the creation of handbooks, such as that of the CUTSD project (Phillips et al., 2000). This resource also incorporates design and implementation guidelines, and focuses on one particular approach (action inquiry). However, unlike the Flashlight tool, no claims are made to general applicability. The handbook clearly contains information and advice that would be of use to any practitioner engaged in action research. However, the fact that it has been created for one specific project has allowed it to provide depth and specialist information rather than the breadth required for a more general resource.

As with methodologies, the range of tools designed to support practitioners is broad, and each has been designed with specific aims in mind. Importantly, little research has been carried out that investigates the impact of such resources. Whilst such tools clearly have the potential to provide great support for practitioners, their relative merits have yet to be fully understood.

## **Authenticity**

Learning technology is, at heart, a pragmatic discipline. A near-universal aim for researchers in this field is to make a difference for students, lecturers or the organisation in which they are based. As a consequence, it is no surprise that the authenticity of evaluation has been of great concern to evaluators working in this area.

Authenticity has been defined as “the notion of how closely an evaluation captures the context of an existing course” (Oliver and Conole, 1998a). Controlled experiments, for example, have been criticised for having little or no relevance to real classroom practice (e.g. Draper, 1997). Moreover, it has been argued that rather than discounting the differences between individuals in order to argue that the sample population is representative, evaluations ought to recognise and understand these differences (Gunn, 1997). Such approaches represents an

important move away from “technological determinism” (Jones, 1999) and towards a situated model of learning (e.g. Brown et al., 1989), reminiscent of the shift from individual to social models of learning in mainstream evaluation.

The result of this position is that methods of evaluation have been developed that concentrate on the notion of authenticity. The approach adopted by the U.K.’s Open University, for example, uses controlled observations for formative evaluation during the development of resources, but supplements this with surveys carried out with real students using the resources as part of their course. Integrative evaluation, as developed by the TILT project (Draper, 1997), focuses on improving the use of resources that have been adopted in a course. Because it is designed for use with real courses where the decision to use learning technology has already been made, it eschews any attempts to make controlled studies of learning technology use. However, it does use quantitative approaches to analyse within-group differences between students. At a greater extreme, ethnography has been advocated as a way of making sense of what really happens on courses where learning technology is introduced, a step that makes authenticity the *sine qua non* of the entire process (Jones, 1998).

## Costs

The evaluation of the costs of learning technology is a highly problematic area, on which comparatively little has been written. However, with the increased drive for accountability in Higher Education, the profile of this issue has been raised considerably.

The area is made complex by a number of issues that remain open for debate. These include (Oliver & Conole, 1998b):

- The meaning of efficiency. (Should it be measured in terms of educational improvement, or cost per day per participant, for example?)
- The identification of hidden costs. (Insurance, travel costs, etc.)
- The relationship between costs and budgets. (Are costs lowered, or simply met by a different group of stakeholders, such as the students?)
- Intangible costs and benefits. (Including issues of quality, innovation and expertise gained.)
- Opportunity costs. (What alternatives could have been implemented? Moreover, if it is problematic costing real scenarios, can costs be identified for hypothetical scenarios and be used meaningfully as the basis for comparison?)
- The use of ‘hours’ as currency. (Are hours all worth the same amount? If salary is used to cost time, how much is student time worth?)
- Whether something as complex as educational innovation can be meaningfully compared on the basis of single figures at the bottom of a balance sheet.

Responses to these issues vary considerably. On one hand, more rigorous methods of accounting for costs have been developed (Bacsich et al., 1999). These have attempted to identify costs that are often neglected and incorporate these into the accounting exercise.

On the other hand, methods have been developed that recognise the problems of costing intangible benefits, for example, and hence use triangulated qualitative data to supplement quantitative information. When used as the basis for decision making, such an approach can have considerable utility, and avoids the difficulties of reducing complex situations to figures (Oliver et al., 1999a). Other alternatives may be feasible under specific conditions. It may be possible, for example, to evaluate changes that cost nothing, or which improve the benefit to cost ratio in a simple way, say through the provision of a new stand-alone resource for students (Doughty, 1998). It is interesting to note, however, that any attempt to account for the complexity of these situations by measuring such benefits soon come to resemble evaluations of educational effectiveness. It may well prove that the most useful way to evaluate such costs is simply to provide an effective method of auditing costs as a complement to mainstream evaluations of learning technology use.

## Checklists

A common method of evaluation advocated in the area of learning technology is the checklist. Essentially, these consist of a list of issues that commonly arise in the process of designing and implementing learning technology. Users are asked to review resources against them and make their judgement based on this structured

consideration. This approach supports decision making by providing a structured approach, and can help prompt potential users to consider factors that they might otherwise neglect.

A wide range of checklists has been developed, each with a slightly different focus. That of Blease (1988), for example, involves an initial categorisation of a piece of software into one of five types (drill-and-practice, arcade, simulation games, lab simulations and content-free tools). Each of these is then supported by an extensive series of questions, which are intended to promote reflection about the appropriateness of the software. These questions are grouped together under headings such as documentation, presentation and layout, friendliness and flexibility, achievement of stated aims, and robustness. Within each section, a range of pragmatic and pedagogic issues are covered, ranging from, “does the program have any accompanying documentation?”, to , “is the software flexible enough to be used in a variety of teaching/learning situations?” Although these questions may well have been useful, the framework cannot claim to be exhaustive (where, for example, would content-based reference materials such as an encyclopaedia be addressed?), and it has also dated. Technological developments mean that many issues are neglected, such as the use of sound and video. In addition, more recent developments, such as the web, are not addressed by the checklist.

Another useful illustration of checklists is provided by the staff development pack developed by the Evaluation and Development Review Unit (EDRU, 1992). These resources are intended to allow practitioners lacking experience in evaluation to assess the impact of a project or activity. The EDRU material comes in two parts: a booklet containing overviews of different aspects of evaluation (EDRU, 1992) and a guide to carrying them out (Sommerlad, 1992). The former contains short papers on a wide range of issues, including curriculum development, staff development, evaluating organisational issues and designing and implementing an evaluation strategy. The latter aims to encourage a ‘process of reflection’ rather than act as a ‘cook book’ for evaluations. It provides a step-by-step guide to evaluation design, starting by assessing the need for evaluation, identifying stakeholders, choosing a method, acting on findings, the role of the evaluator, and so on. Detail is kept to a minimum in order to provide a useful overview. However, this does prevent analytical methods from being covered in any depth, which will cause problems for less experienced evaluators.

Generally, the value of checklists has been called into doubt. A wide range of criticisms has been levelled at them, including (Tergan, 1998):

- A lack of reliability. (Correlation of ratings between reviewers is usually low.)
- Concerns about validity. (Criteria are often based solely on the consensus of developers.)
- Bias. (Themes important to one group of stakeholders, such as designers, may be emphasised at the expense of themes important to other groups, such as students.)
- The aggregation of scores. (The weighting of categories of questions is typically neglected in scored systems.)
- Neglect of individual differences. (Judgement is based on the reviewer’s opinion, rather than on a cross-section of students’ opinions.)
- Neglect of context. (Many reviews focus on the software as an object, and not on its use in a particular context.)
- The lack of tailored criteria. (The categories developed in one context may not be relevant in another.)

A further problem is that whilst checklists help practitioners to ask important questions and identify key issues, they do not help them to resolve these problems. The same lack of expertise that makes checklists useful also makes the more demanding questions difficult or impossible to answer.

Essentially, although such checklists may be useful as prompts for reviewers, or when gathering factual information such as cost or technical requirements, their value as the sole basis for evaluation is limited. This is particularly true when interpretative issues are considered. However, they may be valuable in terms of identifying standards for designers, as discussed below. Another suggestion is that they can be extremely valuable when used alongside other evaluation methods, such as for drawing up a shortlist of packages to be evaluated in greater depth in a second round of evaluation (Le Voi & Morris, 1998). Even when used appropriately, however, it is important that these resources be updated regularly in order to address technological developments.

## **Quality and evaluation**

The introduction of the quality assurance agenda in Higher Education has had many implications for evaluation, both positive and negative. Although it places critical reflection high on institutional agendas, it also fosters a

climate in which negative outcomes are problematic and risks are likely to be replaced by attempts at compliance (Thorpe, 1995).

It is possible that these problems could be avoided by a focus on quality enhancement, rather than on absolute judgements. Such an approach implicitly acknowledges that failures occur, and that there will be, almost inevitably, room for improvement within any educational endeavour. This focus would be well served by ongoing formative evaluation, or by methodologies such as integrative evaluation that focus on improving rather than testing situations.

As well as these concerns, a number of other links between quality and evaluation are worth highlighting. One such link arises from evaluation's role in identifying and documenting examples of good practice (Whittington, 2000). Similarly, it has been argued that checklists have a valuable role to play as standards, for example in the selection of material for inclusion in web-based portal sites (Belcher et al., 2000).

In general, it would appear that evaluation has much to offer to investigators of the quality of learning technology. Given that this area is still being developed, predetermined measures of performance are unlikely to be of great use (Blalock, 1999); consequently, 'open-ended' quality frameworks that use formative evaluation methods to identify the potential for improvement may well be more appropriate (e.g. Oliver et al., 1999b).

### **Changes in methodology through learning technology**

The final theme in this section is the impact that learning technology has had on the process of evaluation. The use of computers to automate routine tasks, such as some types of data collection and analysis, is obvious. However, a number of more subtle changes have also taken place.

This can be illustrated by the example of the changes that arise when a focus group methodology is applied in an online environment. Cousin & Deepwell (1998) observed that this shift resulted in a loss of spontaneity, the inability to take body language into consideration and reliance on participant motivation to engage with the 'pull' technology of bulletin boards (as opposed to 'push' of face to face). On a more positive note, they also noted that this change facilitated data capture, allowed larger groups to participate and gave time for reflection and deliberation.

Such changes are likely to vary from methodology to methodology, and even from situation to situation. Although they may not be predictable, it is important to be aware that tried and tested approaches can give rise to problems in such novel situations. A cautionary tale is given by Jones (1998), concerning the use of ethnographic methods to evaluate online communities. One study saw a group of students "fake" their online collaboration by working together face-to-face to create a script that they could then use to re-create enough online evidence to guarantee them a good mark.

Such considerations highlight the importance of creating a reflective community of practitioners concerned with the evaluation of learning technology. If such issues are to be identified and understood, then reflecting on and sharing experiences of evaluation will be of considerable value to the wider learning technology research community.

### **Conclusions**

Although the notion of evaluation is rooted in a relatively simple concept, the process of judging the value of learning technology is complex and challenging. However, the emergence of a critical literature that reflects on these issues has allowed a number of themes to be identified, and clear developments to be made. These echo discussions that have taken place in the wider evaluation community.

These developments can be broadly characterised by a move from fixed epistemological positions to a situation in which a greater plurality of viewpoints is valued. At the heart of this is a pragmatic focus on the utility of evaluation. Also important in the context of learning technology is an awareness of the priorities that are setting the evaluation agenda. These include the importance of authenticity, the adoption of socio-cultural models of learning and the prevalence of practitioner-based evaluation.



The implications for further research into evaluation are clear. Whilst the influence of the educational context for this work now seems to be agreed upon, greater attention needs to be paid to the impact that learning technology has on the process of evaluation. Additionally, further information needs to be gathered about the most appropriate approaches to adopt in different situations, including strategies for communicating findings effectively. Costs need to be better understood, as do strategies for using checklists and the links between evaluation and other forms of judgement in Higher Education, such as quality assessment. Most importantly, however, this research agenda must be translated into a community of critical reflection that will support and value an increased understanding of these issues.

For practitioners, the situation is more complex. Although tools are being developed which support novice evaluators, the value of these is yet to be demonstrated. Their relative usefulness and effectiveness needs to be understood, and ways of improving both the resources and the way in which they are used must be identified. Moreover, the current climate for learning technology research and development relies on practitioners to learn and do evaluation as one (often relatively unimportant) task amongst many. If the current expectation that practitioners will carry out their own evaluations continues, then strategies for training and supporting them must be developed and implemented.

Evaluation forms a unique meeting point between policy, theory and practice, and as a consequence, it seems unlikely that its practice will ever be uncontentious. Whilst, in an ideal world, it may be desirable to solve these problems, a more immediate priority is that these issues are well understood so that evaluators can respond to them sensitively and appropriately.

## References

- Bacsich, P., Ash, C., Boniwell, K., Kaplan, L., Mardell, J. & Caven-Atack, A. (1999). *The Costs of Networked Learning*, Sheffield: Sheffield Hallam University.
- Becher, T. (1989). *Academic Tribes and Territories: intellectual enquiry and the culture of disciplines*, Buckingham: The Society for Research into Higher Education & The Open University Press.
- Belcher, M., Place, E. & Conole, G. (2000). Quality assurance in subject gateways: creating high quality portals on the Internet. *Quality Assurance in Education*, 8 (1), 38-47.
- Blalock, A. (1999). Evaluation research and the performance management movement: From estrangement to useful integration? *Evaluation*, 5 (2), 117-149.
- Blease, D. (1988). *Choosing educational software*. In Jones, A. (Ed.) *Computers in Education*, Buckingham: Open University Press, 5-13.
- Bradbeer, J. (1999). *Evaluation*, FDTL Technical Report no. 8, University of Portsmouth.
- Breen, R., Jenkins, A., Lindsay, R., & Smith, P. (1998). Insights Through Triangulation: Combining Research Methods to Enhance the Evaluation of IT Based Learning Methods. In Oliver, M. (Ed.) *Innovation in the Evaluation of Learning Technology*, London: University of North London Press, 151-168.
- Brown, J., Collins, A. & Duguid, P. (1989). Situated Cognition and the Culture of Learning. *Educational Researcher*, 18 (1), 32-42.
- Cousin G, & Deepwell, F. (1998). Virtual Focus Group Techniques in the Evaluation of an Electronic Learning Environment. In Oliver, M. (Ed.) *ELT 98: Innovation in the Evaluation of Learning Technology Conference Proceedings*, London: University of North London, 4-7, <http://www.unl.ac.uk/tlrc/elt/elt98.pdf>
- Doughty, G. (1998). Evaluating costs and benefits of investments in learning technology for Technology students. In Oliver, M. (Ed.) *Innovation in the Evaluation of Learning Technology*, London: University of North London Press, 211-222.
- Draper, S. (1997). Prospects for summative evaluation of CAL in higher education. *ALT-J*, 5 (1), 33-39.

- EDRU (1992). *A Guide to Local Evaluation*, Report no. HE/L61, London: Evaluation Development and Review Unit, Tavistock Institute of Human Relations, and the Employment Department.
- Ehrmann, S. (1999). Studying teaching, learning and technology: a tool kit from the Flashlight programme. *Active Learning*, 9, 36-39.
- Gilroy, P., Long, P., Rangecroft, M. & Tricker, T. (1999). The Evaluation of Course Quality Through a Service Template. *Evaluation*, 5 (1), 80-91.
- Guba, E. & Lincoln, Y. (1981). *Effective evaluation: Improving the usefulness of evaluation results through responsive and naturalistic approaches*, London: Jossey-Bass.
- Gunn, C. (1998). CAL evaluation: future directions. *ALT-J*, 5 (1), 40-47.
- Gunvaldsen, J. & Karlsen, R. (1999). The Auditor as Evaluator: How to Remain an Influential Force in the Political Landscape. *Evaluation*, 5 (4), 458-467.
- Harvey, J. (1998). *The LTDI Evaluation Cookbook*, Glasgow: Learning Technology Dissemination Initiative.
- Jones, A., Scanlon, E., Tosunoglu, C., Ross, S., Butcher, P., Murphy, P. & Greenberg, J. (1996). Evaluating CAL at the Open University: 15 Years On. *Computers in Education*, 26 (1-3), 5-15.
- Jones, C. (1998). Evaluation using Ethnography: Context, content and collaboration. In Oliver, M. (Ed.) *Innovation in the Evaluation of Learning Technology*, London: University of North London Press, 87-100.
- Jones, C. (1999). Evaluating a collaborative online learning environment. *Active Learning*, 9, 28-36.
- Le Voi, M. & Morris, E. (1998). Systematic Evaluation of Large Corpora of Computer-Based Teaching Resources. In Oliver, M. (Ed.) *ELT 98: Innovation in the Evaluation of Learning Technology Conference Proceedings*, London: University of North London, 15-19,  
<http://www.unl.ac.uk/tltc/elt/elt98.pdf>
- Light, R. & Smith, J. (1970). Choosing a Future: Strategies for Designing and Evaluating New Programs. *Harvard Educational Review*, 40, 1-28.
- Oliver, M. & Bradley, C. (1999). *Examples of Best Practice in the use of multimedia in Higher Education*, EXE report no. 2, London: University of North London.
- Oliver, M. (1999a). Evaluating Learning Technology: issues for practitioners. *Active Learning*, 8, 3-8.
- Oliver, M. (1999b). *The Evaluation of Learning Technology Conference Proceedings*, B. P. ELT Report, London: University of North London.
- Oliver, M. (1999c). *The B.P. Evaluation of Learning Technology Project: a review*, B.P. ELT Report no. 9, London: University of North London.
- Oliver, M. (1999d). *The ELT Toolkit*,  
<http://www.unl.ac.uk/tltc/elt/toolkit.pdf>
- Oliver, M. & Conole, G. (1998a). Evaluating Communication and Information Technologies: A toolkit for practitioners. *Active Learning*, 8, 3-8.
- Oliver, M. & Conole, G. (1998b). The Evaluation of Learning Technology – an overview. In Oliver, M. (Ed.) *Innovation in the Evaluation of Learning Technology*, London: University of North London Press, 5-22.
- Oliver, M. & Conole, G. (1999). A response to Ehrmann. *Active Learning*, 9, 40.
- Oliver, M., Conole, G. & Kewell, B. (1998). An overview of an evaluation toolkit for learning technology use. In Dooner, M. (Ed.) *Proceedings of the IEE Colloquium, "Exploiting Learning Technology: issues for learners and educators"*, London: Institution of Electrical Engineers, 1/1-1/4.

- Oliver, M., Conole, G. & Bonetti, L. (1999a). *The hidden costs of change: Evaluating the impact of moving to online delivery*, TLTC report no. 1, London: University of North London.
- Oliver, M., Phelps, J., Beetham, H., Bailey, P. & Conole, G. (1999b). *Evaluating EFFECTS: identifying issues and assessing the evaluation framework*, EFFECTS report no. 3, London: University of North London.
- Patton, M. (1997) *Utilization-focused evaluation*, London: Sage.
- Phillips, R. Bain, J., McNaught, C., Rice, M. & Tripp D. (2000). *Handbook for Learning-centred Evaluation of Computer-facilitated Learning Projects in Higher Education*, <http://cleo.murdoch.edu.au/projects/cutsd99/handbook/handbook.htm>.
- Preskill, H. & Torres, R. (1999). Building capacity for organizational learning through evaluative enquiry. *Evaluation*, 5 (1), 42-60.
- Schön, D. (1983). *The reflexive practitioner: how professionals think in action*, New York: Basic Books.
- Sommerlad, E. (1992). *A Guide to Local Evaluation*, Report no. HE/L62/1186, London: Evaluation Development and Review Unit, Tavistock Institute of Human Relations, and the Employment Department.
- Tavistock (1999). *Teaching and Learning Technology Programme Guidelines for Project Evaluation*, <http://iet.open.ac.uk/tltp/bs031doc.html>.
- Tergan, S. (1998). Checklists for the Evaluation of Educational Software: Critical Review and Prospects. *Innovations in Education and Training International*, 35 (1), 9-20.
- Thorpe, M. (1995) Evaluating Technology-Based Learning. In Collis, B. & Davies, G. (Eds.) *Innovative Adult Learning with Innovative Technologies*, Amsterdam: Elsevier, 191-199.
- Torres, R., Preskill, H. & Piontek, M. (1996). *Evaluation strategies for communicating and reporting: Enhancing learning in organizations*, London: Sage.
- Whittington, D. (2000). Evaluating three years' use of a virtual university. *Quality Assurance in Education*, 8 (1), 48-52.