# Reflections on a model of evaluating learning technologies

## Book Section

# Innovation in the Evaluation of Learning Technology

Edited by

Martin Oliver
*University of North London*

First published 1998 by Learning and Teaching Innovation and Development, University of North London. Printed and bound in Great Britain by the University of North London, 236-250 Holloway Road, London, N7 6PP.



ISBN 1-85377-256-9

# Chapter 2
# Reflections on a model for evaluating learning technologies

Ann Jones, Eileen Scanlon and Canan Blake

*This chapter discusses an evaluation framework, CIAO!, developed by the Computers and Learning Research Group at the Open University and used to evaluate students' use of technology based teaching and learning on a number of different courses.*

*The framework has been described briefly elsewhere (Jones et. al. 1996) and case studies of its use have also been documented (Scanlon et. al, 1998). This chapter provides a critical reflection on its usefulness for evaluating teaching and learning technologies. The chapter begins with a discussion of important issues in evaluating technology-based teaching and the kinds of information that can be collected in an evaluation. It then introduces the framework and uses one of the evaluation studies conducted at the Open University in order to discuss the advantages and disadvantages of the framework.*

## Introduction

Four years ago the Computers and Learning Research group at the Open University (OU) began a project to evaluate teaching and learning software developed by the OU for our undergraduate students. In order to do this the team developed an evaluation framework, CIAO!, that was informed by extensive experience of evaluation and also by the literature on evaluating technologies for teaching and learning. The framework has now been used to evaluate students' use of technology based teaching and learning on nine different courses.

In the first half of the chapter, the framework is used to discuss issues that need to be considered when evaluating CAL. In the second half of the chapter we use one evaluation case study, of computer conferencing and resource based learning in the Humanities, to reflect on the strengths and weaknesses of the framework.

## Issues in Evaluating Computer Assisted Learning (CAL)

The CIAO! framework, developed to support the evaluation of CAL developments at the Open University (Jones et. al., 1996) outlines three dimensions to evaluation: (i) context, (ii) interactions and (iii) attitudes and outcomes. This is shown in Table 1. This first section discusses the critical issues in evaluating learning technologies that led to the emphasis on these three areas.

Table 1: The CIAO! Framework (Source: Jones et al, 1996, p12)

| | Context | Interactions | Attitudes and outcomes |
|---|---|---|---|
| **Rationale** | The aims and the context of use | Data about students' interactions with the software allows some analysis of the process | It is important to try to assess learning outcomes but also to consider affective outcomes, e.g. perceptions and attitudes |
| **Data** | Designers' and course teams' aims;<br>Policy documents and meeting records | Records of student interactions<br>Products of students' work<br>Student diaries<br>On-line logs | Measures of learning.<br>Changes in students' attitudes and perceptions |
| **Methods** | Interview CAL Designers and course team members.<br>Analyse policy documents | Observation<br>Diaries<br>Video/audio and computer recording | Interviews<br>Questionnaires<br>Tests |

### Context

Context can refer to a number of aspects including the setting the teaching is in (e.g. its location and who is involved), the way a particular application is used, and the designer's goal in introducing the CAL; as such, it is very important in an evaluation. In the framework it encompasses a number of areas including the rationale for introducing the technology, the context within the course (i.e. how it fits with other course components), the context of use (e.g. is it to be used at home, by individuals or groups, or in a classroom) and to what extent and how is it supported by human teaching? Context is not a recent concern: in their illuminative evaluation approach

Parlett and Hamilton (1987) emphasised that "the innovation is not examined in isolation but in the school context or 'learning milieu'. In the recent evaluation literature there is also an increasing concern with context: e.g. Draper et. al. (1996) stress the importance of conducting the evaluation on the target group who will be using the software; Joyce et. al. (1998) emphasise the importance of taking account of the "whole learning experience" and Oliver and Conole (1998) propose three evaluation qualities including *authenticity* which "describes the notion of how closely an evaluation captures the context of an existing course" (Oliver and Conole, 1998).

A crucial aspect of context is the designer's rationale in introducing the technology. Analysis and understanding of the rationale is essential in determining the evaluation questions to be asked. For example, two different rationales for the CAL evaluated in the OU study include:

1. the use of multi-media to bring together resources that would otherwise be separate (e.g. in video, text, audio, photographs) and to make them available – so as to allow the student to take on the role of the researcher

2. the use of CAL to teach concepts known to be difficult, or dynamic methods to illustrate dynamic concepts or functions: in biology, for instance, this might include the respiratory system. The first evaluation undertaken in our study was of a CAL program designed to help students understand Phase Diagrams (Tosunoglu et. al., 1996) which students had previously found difficult.

These different rationales will require different evaluation questions: in the first, they include *Can the student really experience being a researcher? Are the media integrated? Can the student navigate around them?* In the second, the main question is whether this approach (the use of CAL) helps students to learn difficult concepts.

Before we can find out whether the designer's motivation is justified, we need to know what that motivation is and one way is to ask the designers/teachers what benefits they are expecting from the learning technology. We are not suggesting that the evaluation questions are driven only by the teacher or designer's view, but that understanding the designer's aim is a crucial element.

Mason (1996) also emphasises this issue in her proposed approach to evaluation, which includes 'drawing out the central issues' as the first step. She suggests that this requires the evaluator to be "*responsive to the application as a whole, looking at the overall aims as well as the details, seeking the views of all parties involved, from the stakeholders and implementors to the teachers and students" (p4)*

Another important aspect of context is the features of the setting, e.g. is the software intended to stand alone (as much multimedia is now)? If so, it is important to ensure that students can work through it successfully without needing much help. In evaluating students' use of technologies at home, in a distance learning context with little human support, there is therefore much emphasis on the ease of setting up and using the software.

In such contexts, the extent to which students can navigate through the package successfully is also an important issue. Whilst resource based learning offers the potential for a much more student-centred environment, there is increasing awareness of difficulties students may have in using such environments for learning (Taylor and Laurillard, 1995) and of the skills they will need (McDonald and Mason, 1997).

It is also important that the evaluation is 'open' enough to reveal unexpected findings, as well as what is being investigated. This was a key point in Parlett and Hamilton's illuminative evaluation, mentioned earlier. Later, an example is given of a formative evaluation where it became apparent that the guided approach in the software was very helpful for some students and quite problematic for others. If the evaluation had relied on pre-conceived questions, this would not have been revealed. This is a point also stressed by Joyce et al (1998) who advocate a largely qualitative approach to evaluation in order for the researcher to get close to the issues of the evaluation and to have a feel for the context.

**What kind of data should be collected?**

When evaluations are commissioned, clients often start by asking evaluation questions such as: *Do students learn better, faster or more using this technology?* and can be frustrated when evaluators tell them that these are not appropriate questions, and moreover they are not questions that they can answer! The main problem here is that in order to answer them, experimental or quasi-experimental designs are necessary. In a typical design, a group learning through the use of the computer based technology will be compared with a group that is not, by giving each pre and post tests

and comparing the gains made. However, there are problems in applying this kind of approach to education as Gunn (1997) points out and also Draper (1997) (who is focusing on the problems of summative evaluation) and so it is worth summarising some of the main issues:

*1. Each learning context is different and because the learning technology is only one component, it is difficult to make such comparisons*

The experimental approach depends on being able to control all the variables except the one under investigation. In the 'hard' sciences this can be done relatively easily under laboratory conditions: e.g. we can vary light levels to investigate the effects on plant growth, keeping other variables constant. However, it is much more problematic in educational settings.

*2. Where there is appropriate use of learning technologies, it will often change the learning experience*

It has always been the case that computers change what is learnt, so that comparisons are problematic. However, the more recent combination of developments in communication technologies and their widespread availability, and the computer's capabilities for manipulating and presenting information offers the potential for transforming teaching and learning. Such changes include a shift to collaborative learning (from more individually based learning); a shift to resource based learning which is much less teacher directed and this inevitably leads to a change in the distribution of expertise and in the roles of both teacher and learner.

In discussing learning science on-line Scanlon (1997) describes how the development of communications technologies is redefining science and in turn, redefining what it means to learn science. One example she gives is of a project that uses scientists as experts offering advice to young children. The project:

> *uses a network connection to foster a collaboration between experts at the San Francisco Exploratorium, a class of ten year olds and their teacher. ...(Students can)... create messages containing pictures and text as well as video. Students... chose a weather phenomenon to study in depth and made video observations. Questions about these phenomena... were sent to the experts who replied with video annotations. Students then helped each other understand the experts' messages and collaborated on new messages and questions."*

*(Scanlon, 1997, p65)*

*3. The timescale is problematic: when do you do the testing?*

When using pre- and post tests, as the terms imply, measures of learning are taken before and after the students' use of the learning technology. But when exactly should the post test be given? Tests given to students just after they have used the package may indicate their mastery at this point, but won't reveal whether the material is retained. An argument can therefore be made (see Issroff (1995) and Draper et. al., 1996) that delayed post tests, given some weeks later, can be helpful in picking up any longer term effects, but as Draper et al. point out, changes may then be due to other factors, such as revision classes.

The problem in deciding when to administer post tests is part of the bigger issue discussed earlier, of applying the experimental model to an educational context, where it is very difficult to control all the variables.

Having made all these qualifications, there are times when it can be appropriate to make such comparisons and use pre and post tests. Comparisons between two versions of learning technology or between technology and non technology based teaching make most sense in two particular contexts:

- when we are interested in which of two possible approaches in learning technologies might be better
- where one media is replacing another

For example, one course evaluated in our project had remedial numeracy materials in two forms: as CAL tutorials and as written booklets. Here there was an opportunity to compare these media by collecting data on students' usage and their attitudes. The vast majority of the students stated that they preferred having both media (Jones et al, 1998). Comparative measures are therefore included in the CIAO! framework, along with other measures. This approach, of using a range of different methods is increasingly used in educational evaluation (see e.g. Draper et al, 1996; Milne and Heath, 1998; Joyce et. al., 1998).

**Collecting attitudinal data**

So far the discussion has focused on cognitive benefits: e.g. measuring changes in students' learning. It is equally important to collect information about students' perceptions and attitudes.

Draper et. al. (1996) argue that attitude data is problematic in evaluating CAL: attitudes to CAL vary widely within and across classes and these positive and negative views may be unrelated to the educational value of the software. They also suggest that students' expectations of CAL differ considerably to their expectations of other forms of teaching, e.g. lectures. We have collected attitude data in distance learning over a number of years and have found more consistency in students' attitudes than is suggested by Draper although we would agree that they are weak measures of educational effects. However, there is increasing recognition of the need to pay closer attention to the social and psychological factors in technology based learning and increasing evidence of these factors mediating learning.

For example, a study by Issroff (1995) compared the performance of individuals and pairs in the collaborative use of CAL in Chemistry. As in many other evaluation studies, there were no differences between the individuals and pairs when learning outcomes were measured by post tests. However, the perceptions and motivations of the students in the two conditions differed: the students who worked in pairs were more positive, enjoyed the experience more and were more motivated. There were also differences in the work produced by the students whilst they used the computer (in this case worksheets) and social factors affected the nature of the interaction. Studies such as this confirm the importance of paying attention to affective aspects as well as cognitive factors, and of using information from a variety of sources, including, where appropriate, pre- and post-achievement tests, interviews and questionnaires with students and teachers to assess attitudes and perceptions.

Our framework includes both interviews and questionnaires, as these serve different functions. Whereas through questionnaires, learners can only respond to the questions asked, rather than the issues they would like to raise, interviews allow a two way dialogue. Such dialogues are essential if the evaluation is to be open to the issues that are of concern to the learners.

It can be helpful to use both questionnaires and interviews at different points in the evaluation process. For example, interviews to explore the issues, followed by a questionnaire which surveys these particular issues, or

a questionnaire followed by interviews to allow exploration of the issues and to move into other areas.

**Collecting information about the process: tracking interactions**

The kind of data we have discussed so far, where we are essentially investigating what has changed as a result of using the learning technology, be they changes in perceptions, attitudes or motivation, are referred to as outcome data in our CIAO! framework. Outcome data alone are very limiting because it can only ever tell us that something has occurred, not why and how. However, evaluation is usually intended not just to provide information about whether an innovation has been successful but also so that improvements can be made (see, e.g. Calder, 1994). Hence, in addition to any outcome data it is at least as important (and some would argue more important) to try to understand the process.

One way of beginning to understand this process is to look in detail at students' use of the software by analysing their interactions with the software. This can provide information on why and how particular elements work (or do not) rather than just finding out what does and doesn't work. Where possible, students are observed working with the software: for us in a distance learning context this may be at residential schools; in their homes or they may be invited to come to the campus, as in the case of the Living with Technology course which is discussed below. Such interactions can be audio or video-taped, in order to provide protocol data for later analysis. Computer logs can also be collected of all key presses and the routes that students take through the materials. There may also be products of student' work (e.g. worksheets completed whilst using the package).

Collecting process data can be particularly important and illuminating where the purpose of the evaluation is formative, i.e. it is providing feedback that contributes to the development of the package. For example, in one formative evaluation study of the Living with Technology course (Jones et al, 1998), we investigated students' use of a CAL package that aims to teach precedence rules in manipulating algebraic equations to students on a foundation technology course. Local students were invited to the university campus for an evening session in which they used the software and members of the evaluation team observed and noted any problems they had and any unexpected events. At the end, students' comments were gathered from a discussion and via a questionnaire.

An unexpected issue emerged about how best to support procedural skills. The program starts off being quite directive and taking students through each step of the calculations then it requires students work out some of the calculations themselves. Some students were uneasy with this guided approach: they felt that they were working at a mechanical, surface level without having to understand the calculations. Other students felt they were able to get an overall view of the problem without getting lost in the detail.

Such problems and issues, which led to changes in the software, could not have been revealed without such observations. This observation study is an example of a formative evaluation where educational software is evaluated whilst still under development in order to inform changes to the software. We have found such formative evaluations to be extremely important and are a routine part of our course development where learning technologies are used.

In such formative evaluations an iterative model is often used (see Jones, Kirkup and Kirkwood, 1993) which includes different evaluation phases. In a typical example, the first phase will be to elicit critical comments from external experts. Students will then work through the whole course in detail, the course will be revised following their feedback and finally the revised version will be tested out by a small group of students. This cyclical process shares many aspects of the action research model described by Lousberg and Soler (1998). However, in our case it has a limited number of cycles, and in formative evaluations stops when the course is complete, although such ongoing research is conducted into students' use, ownership and attitudes towards new technologies.

**A critical review of CIAO!'s use in evaluating technology based learning in a social history course**

*The course and the rationale for the use of CAL*

*Charles Booth and Social Investigation in Nineteenth Century Britain* is an advanced undergraduate course. Students use the computer as a tool for historical research for constructing historical databases, and to help formulate their own research question. Students work from home using their computer and the written study guide and course books. As the course has a very low population that is geographically dispersed, there are no face-to-face tutorials, so all tuition is via the computer conferencing environment, except for one day school at the OU campus.

All the original and essential data for the course is on the CD-ROM which provides a multimedia data resource. It contains primary data from diverse media including published writings, transcribed manuscript material, Booth's colour coded maps of Victorian London, O/S maps, illustrations and plans, original census data, photographs, audio, biographies, bibliographies, glossaries and on-screen tutorials. One aim of the computer conferencing environment is to provide teaching that supports the CD-ROM resource.

*The questions or issues under investigation*

Students needed to acquire a reasonable level of computer literacy in order to make effective use of the CD-ROM, which was essential for their course work as well as computer conferencing. Students were advised that some experience in using PCs was desirable. However a range of expertise and experience was likely and as students were working on their own, they needed to be able to use the technology with minimum help. It was important to find out, therefore, whether students could cope with the technology; whether the on-screen tutorials succeeded in guiding students around the CD-ROM, and the contribution of the conferencing system. The evaluation questions were:

- What problems do students face in using the CD-ROM for their coursework?
- How much do they use the on-screen tutorials and how useful do they find them?
- How much was the FirstClass conferencing system used by students and what was their perception about its contribution to their studies?

*Choice of methods*

There was a limited opportunity to observe students working at a day school and also to interview four students. Outcome data was collected by a questionnaire. There was also a computer conference for problems related to the CD-ROM as well as the on-line tutorials that provided further interaction data. This is summarised in Table 2 below:

Table 2. Framework for evaluating the 'Charles Booth' course
(Source, Jones et al, 1998)

| Context | Interactions | Attitudes and Outcomes |
|---|---|---|

| | | | |
|---|---|---|---|
| **Methods** | Two uses of TLTs: computer conferencing and resource based learning (RBL).<br><br>CD-ROM for RBL for student to acquire computer skills for research/research skills.<br><br>Computer conferencing replaces F2F tutorials | Students were observed constructing historical databases at a day school.<br><br>Four students were interviewed.<br>All computer conference messages were collected | Questionnaires designed to investigate students' use of the CD-ROM and also their attitudes and perceptions.<br><br>A computer conference for problems related to the CD-ROM. |
| **Data** | Course descriptions, article on course.<br><br>Video interview for course team chair on rationale.<br><br>Discussions with course team chair. | Notes of students' facility with and problems in constructing databases.<br>Interview transcripts<br>Messages from the computer conferencing tutorials | Students' use, attitudes and perceptions.<br><br>Students' use of the CD-ROM. |

*Brief summary of findings*

(See Tosunoglu-Blake et. al., 1998, for full report)

All students were able to install and set up the software successfully. As in other courses using CD-ROMs (MacDonald and Mason, 1997) students found that reading the CD-ROM on the screen was tiring and so they printed some of the material and also requested that the instructions be in printed form.

Students did not report any serious navigation problems in accessing the different parts of the CD-ROM. Most students praised the CD-ROM ('it brought (the) archives into immediate personal contact') although some students found it time consuming because of the richness of the materials and the subsequent side tracking. They were also positive about facilities such as the search facility that could be used to analyse Booth's texts, the databases and the census data.

The computer conferencing included structured academic conferences for each of the four tutorial groups, along with social and 'chat' conferences and technical help. The feedback on these was very positive. Whilst the majority of students used the conferences to read other messages and learn from messages that applied to their situation, nearly as many also contributed to academic conferences, and two thirds used it to ask their tutor for help.

Apart from helping them to feel less isolated, students also commented that the process of writing the messages helped to clarify their thoughts. The students' contributions showed that they used the conferencing to propose and develop academic arguments about social investigation which made use of the CD-ROM, and which acknowledged and built on other students' contributions. Below is an extract from a student's contribution following a tutor's question about the difference between sociology and social investigation:

> *I support Megan's comments but with a little reservation as some sociological studies are conducted in great detail (e.g. Goffman's work) whilst some sociological investigations are limited in scope and size*

*Commentary*

The evaluation showed that the CD-ROM did indeed allow distance learners to obtain primary data from diverse resources in order to carry out project work. The combination of a small amount of observation and interviewing, together with the data from the questionnaires allowed judgements about the usability of the CD-ROM software and revealed the students' perceptions. Access to the conferences enabled an evaluation of both the amount of contribution made and the nature of the debate.

The methods used, summarised in table 2, allowed us to answer the questions posed earlier. The framework is not particularly geared towards communications technologies and does not, therefore, suggest any particular methods for analysing students' conference contributions. We have already argued for the need to analyse interactions to understand processes: we also need to analyse conference contributions in order to understand the use of conferencing. In a continuation of the evaluation study, we therefore investigated the patterns of participation, the content of the messages and the extent to which they focus on the academic debate. These were investigated within the context of a particular academic conference.

Contributions were analysed in two different ways. The number of messages and their length were recorded and also whether they were from students or tutors in order to investigate the pattern of participation. There is plenty of literature documenting such methods and it is helpful in investigating the patterns of participation, but we would agree with Mason (1996) that:

> *simplistic counting of entries without any regard for the content of the message does nothing to progress the use of computer conferencing as an interactive educational medium.... (p5)*

We also wanted to scrutinize the content of the messages: here, our interest was in investigating how students went about answering the questions posed and the extent to which they referred to other messages. Our approach, therefore, was to use standard qualitative techniques of discourse analysis: deriving categories using essentially a grounded theory approach (Strauss, 1987). The categories derived from it so far have provided helpful evidence about the nature of the conference contributions and the extent to which they contribute to an academic debate (Jones, Blake and O'Day, 1998).

**Possible disadvantages of the CIAO! Framework**

In the same way that we have argued that the context of an evaluation is crucial, the context in which a framework such as CIAO! is developed is also significant: in the case of the CIAO! framework there are specific features about our students and their location, and our infrastructure, that may have implications for generalisability.

*Access to CAL developers/course team*

CIAO! was developed in the context of a team approach and has mostly applied to software developed at the Open University and so we have been able to ask developers and course chairs about their rationale. Clearly this would be harder when evaluating software developed elsewhere, yet we would argue that it would be important to understand the rationale in order to judge its suitability for use in a different context as well as for its evaluation.

*Large scale use*

It has been suggested (Oliver and Conole, op. cit.) that the framework has a focus on survey methods and so is most suitable for large-scale use: i.e.

where there is a large student population. It is certainly true that many courses do have large populations, and it is also the case that we have the support of a dedicated survey office for sending out and analysing questionnaires, which eases the work of the evaluators. However, the university also has a commitment to small-scale courses (such as in the case study above) and it is particularly important to track the exploration of small scale innovations, which may eventually scale up. In these situations, methods other than postal surveys are likely to be prominent.

*Student characteristics*

The Open University routinely surveys students as part of quality assurance processes, and we are fortunate that we usually have a high response rate. In contexts where the response rate is likely to be significantly lower evaluators will need to consider other ways to ensure the data is reasonably representative. However, although we may be fortunate in our response rates, carrying out observational studies is particularly tricky in our context where our options are to observe students in their homes; at residential schools (if these exist) or to bring students on to the campus. We envisage that observation should be easier in universities where students are on campus.

Finally, it is clear that the boundaries between Open and Distance Learning and conventional on-campus universities are continually blurring and changing. Two consequences of this are first of all an increasing need to understand the context of technology based learning and take it into account and second that the two communities should increasingly be able to share and benefit from each other's practices and experiences.

**Conclusions**

The aim of the CIAO! framework is to provide a variety of methods to be considered for particular evaluations, rather than, for example, to be prescriptive in suggesting a particular approach. Its overall focus is educational, with at least as much emphasis on process as outcomes, which in turn leads to an emphasis on using observation to collect empirical data. Experience suggests that a choice of methods rather than prescriptions about particular methods or checklists means that it can be usefully applied in a wide range of contexts. Other approaches to evaluation also emphasise the need for a range of tools and the problems of controlled studies. However, the lack of prescription in the CIAO! framework means that there is no detailed guidance on how to apply particular methods to contexts where

they may not traditionally have been used. We do not think this is problematic: good evaluation will involve an understanding of the range, appropriateness and applicability of methods, as with any research skills: evaluation is not a special case. It does mean, however, that it will be important for the evaluator to have an understanding of learning technologies: in particular to appreciate what benefits particular technologies may offer. We believe that using learner centred evaluation methods will help students to make the best use of the technologies on offer.

**Acknowledgements**

We would like to thank our many colleagues who contributed to the various case studies reported here: Rosemary O'Day; David Englander; Tony Anthoni; Peter Cox; Karen Shipp; Tony Nixon; Dick Morris; the associate lecturers who took part in the day schools, and, of course, the students.

**References**

Calder, J (1994) *Programme Evaluation and Quality: A Comprehensive Guide to Setting up an Evaluation System*, Kogan Page.

Draper. S. W., Brown, M. I., Henderson, F. P & McAteer, E. (1996) 'Integrative evaluation: an emerging role for classroom studies of CAL', *Computers and Education*, Vol. 26, No. 1-3, pp 17-32.

Draper, S. W. (1997), 'Prospects for Summative Evaluation of CAL in Higher Education', *ALT-J*, Vol. 5, No. 1, pp 33-39

Gunn, C (1997) CAL evaluation: future directions, *ALT-J*, Vol. 5, No. 1, pp 40-47

Issroff, K, Scanlon, E & Jones, A. (1995) Methodological Considerations for Investigating Computer-supported Collaborative Learning. *Proceedings of WICCE '95, Birmingham, UK.*

Jones, A. Kirkup, G & Kirkwood, A. (1993) *Personal Computers for Distance Learning.* Paul Chapman.

Jones, A., Scanlon, E., Tosunoglu, C., Ross, S., Butcher, Murphy, P. & Greenberg, J. (1996) Evaluating CAL at the Open University: 15 years on. *Computers and Education*, Vol. 26, No. 1-3, pp 5-15.

Jones, A., Scanlon, E., Tosunoglu, C., Butcher, Greenberg, J. & Ross, S. (1998) Contexts for evaluating educational software. *Interacting with Computers*, In Press.

Jones, A., Tosunoglu-Blake, C, & O'Day, R. (1998) *Computer Conferencing in A427: A Preliminary Report*, CAL Research Group Technical Report, June, Insitute of Educational Technology, Open University.

Joyce , G., Choo, B. G., Pavix, A., Newman, J., Beeby, A. and Wainwright, P. (1998) Learning outcomes and the role of evaluation in Embedding COMPACT. *Active Learning* 8, July, pp32-35.

Lousberg, M & Soler, J. (1998) Action research and the evaluation of IT projects. *Active Learning* 8, July, pp36-40.

Mason, R. (1995) Evaluating Technology-Based Learning, in B. Collis and G. Davies, (eds.) *Innovating Adult Learning with Innovative Technologies*. Elsevier, Amsterdam.

MacDonald, J & Mason, R. (1997) Information Handling Skills and Resource Based Learning. *Open Learning*, Vol. 12, No. 3, In Press

Milne, J & Heath, S (1998) *Evaluation Handbook for successful CAL courseware development*. Centre for CBL in Land Use and Environmental Sciences, Aberdeen University

Oliver, M & Conole, G (1998) Evaluating communication and information technologies: a toolkit for practitioners. *Active Learning* 8, July, pp3-8.

Parlett, M. & Hamilton, D (1987) Evaluation as illumination: a new approach to the study of innovatory programmes. In Murphy, R. and Torrance H., *Evaluating education: issues and methods.* London: Harper and Row Ltd.

Scanlon, E. (1997) Learning Science On-line. *Studies in Science Education, 30 ,* 57-92

Scanlon, E., Tosunoglu, C., Jones, A., Butcher, B., Ross, S., Greenberg, J., Taylor J. & Murphy, P. (1998) Learning with computers: experiences from an evaluation project. *Computers and Education*, Vol. 30, 1/2, pp9-14

Scanlon, E., Ross, S. & Tosunoglu, C. (1997) Learning with computers: experiences from an evaluation project. *Computers and Education*, Vol. 30, 1/2, pp9-14

Strauss, A. L. (1987) *Qualitative Analysis for Social Scientists*, Cambridge University Press

Taylor, J. & Laurillard, D. (1995) Supporting Resource Based Learning. In N. Heap, R. Thomas, G. Einon, R. Mason and H. Mackay, (eds.) *Information Technology and Society*, Sage.

Tosunoglu, C., Butcher, P., Scanlon, E & A. Jones., (1996) The Works Metallurgist: an evaluation of a CAL package on phase diagrams. *ALT-J,* Vol.4, No.1, pp 51-57.

Tosunoglu Blake, C., Jones, A., O'Day, R., Englander, D & Cox, P. (1998) *Virtual Archive - an evaluation*, CALRG Report, Institute of Educational Technology, Open University, June.