CAL evaluation:

future directions

Cathy Gunn

Centre for Professional Development, University of Auckland, New Zealand

Formal, experimental methods have proved increasingly difficult to implement, and lack the capacity to generate detailed results when evaluating the impact of CAL on teaching and learning. The rigid nature of experimental design restricts the scope of investigations and the conditions in which studies can be conducted. It has also consistently failed to account for all influences on learning. In innovative CAL environments, practical and theoretical development depends on the ability fully to investigate the wide range of such influences. Over the past five years, a customizable evaluation framework has been developed specifically for CAL research. The conceptual approach is defined as Situated Evaluation of CAL (SECAL), and the primary focus is on quality of learning outcomes. Two important principles underpin this development. First, the widely accepted need to evaluate in authentic contexts includes examination of the combined effects of CAL with other resources and influential aspects of the learning environment. Secondly, evaluation design is based on a critical approach and qualitative, case-based research. Positive outcomes from applications of SECAL include the easy satisfaction of practical and situation-specific requirements and the relatively low cost of evaluation studies. Although there is little scope to produce generalizable results in the short term, the difficulty of doing so in experimental studies suggests that this objective is difficult to achieve in educational research. A more realistic, longer-term aim is the development of grounded theory based on common findings from individual cases.

Experiments that failed

Scientific, experimental methodology was previously considered to be the only acceptable approach to educational research. Two important principles of experimental design are:

- to balance individual differences within study populations and so achieve generalizable results,

and

- to attempt to isolate the effects of a single resource for evaluation purposes.

Problems with this approach were reported in the literature of the 1970s (Elton and Laurillard, 1979; MacDonald and Jenkins, 1979) when the influence on learning of

individual and contextual factors was recognized. Similar issues emerged during the 1980s and early 1990s, (Bates, 1981; Spencer, 1991) when the inability to identify which single or combined factors supported learning became a recurrent problem. It was clear that prior knowledge, approaches to learning, provision of appropriate scaffolding, complementary combinations of resources and various contextual factors all influenced the quality of learning outcomes. It was concluded that evaluations must be designed to account for these factors, rather than to balance or disregard them as was previously the norm (Kemmis, 1987, Gunn, 1995).

Another problem stemmed from the belief that single studies involving large sample sizes were necessary to produce meaningful results. The rather indiscriminate choice of study populations required to produce the requisite numbers frequently resulted in low motivation and levels of perceived relevance of evaluation tasks to personal and educational goals (Draper *et al*, 1996; Gunn, 1996). This suggested that the true potential for learning with CAL could not be reliably assessed unless its use formed an integral part of a course, and evaluations involved only the students on that course. It was thus concluded that the more specific aspects of CAL evaluation could not be served by a general and inflexible research methodology originally designed to measure the uniform and largely predictable behaviour of organisms in the physical sciences.

The basis of an alternative methodology

In the context of the work reported here, development of a suitable methodology began with a review of educational research literature. Critical theory (Carr and Kemmis, 1986), critical ethnography (Angus, 1986) and qualitative methodology (Denzin and Lincoln, 1994) were adopted as the basis for grounded development of the Situated Evaluation of CAL (SECAL) framework. The stages of development are described elsewhere (Gunn, 1995; 1996). Within the situated approach, the standard range of objective and subjective research methods are used, as appropriate, for evaluation study design, i.e. observation, field notes, log data, interview, questionnaires, outcomes analysis, results comparisons, attitude surveys, expert review, and discussion groups.

The SECAL approach is opportunistic in recognizing that situation-specific factors such as logistical and ethical constraints will determine what subset of qualitative methods is available, and of these methods, which are the most appropriate in a particular case. The case-study method described by Yin (1991) is the basis for study design, and from this comes the longer-term objective to develop grounded theory (Glaser and Strauss, 1967). Action research (Zuber-Skerrit, 1990) describes the preferred, collaborative approach involving all interested parties in CAL development and integration initiatives. It also defines the action-reflection-modification cycle that models the dynamic process of educational change.

The concept of situated evaluation

The concept of situation is described in relation to knowledge and learning by Brown *et al* (1989): 'Knowledge is situated being in part a product of the activity, context and culture in which it is developed and used'. Although there is some debate about the precise nature

of situatedness, if the concept is accepted at all, evaluation should not assess whether CAL alone supports learning or enhances its effectiveness. It must examine the effects of particular resource combinations within specific learning contexts. It was never considered necessary to evaluate the impact of a book or lecture in isolation, so the justification for doing so with CAL is hard to see.

The range of potential influences on learning is broad, and includes factors which are both intrinsic and extrinsic to any particular study resource. To accommodate all possibilities, SECAL includes the concepts of *evaluation in context* and *evaluation of context*. *Evaluation in context* refers to study of the primary effects of using CAL programs with other resources and forms of support. The proper integration of CAL into courses is crucial. *Evaluation of context* examines factors related indirectly to a CAL program or the immediate learning environment, but ones which can still influence integration at an institutional level, and so impact on learning outcomes. Factors related to levels of institutional support for acquisition, development and use of CAL fall into this latter category. There is also a dynamic aspect of SECAL which supports recommendation and implementation of beneficial changes to learning environments.

Authenticity and context

The importance of contextual influences implies that authenticity in study design should be a non-negotiable factor. Evaluation is consequently limited in scope and frequency by the number of target users and available opportunities for the effects of CAL to be evaluated as a fully integrated part of a course. Although these limitations may appear to be rather restrictive, the comprehensive and theoretically supported SECAL framework has been successfully applied in very different circumstances and has produced relevant and meaningful results. The experimental objectives of generalizable results and theory generation are not ruled out – they just take longer to achieve. In view of the strict relevance of outcomes, the compromise is worthwhile.

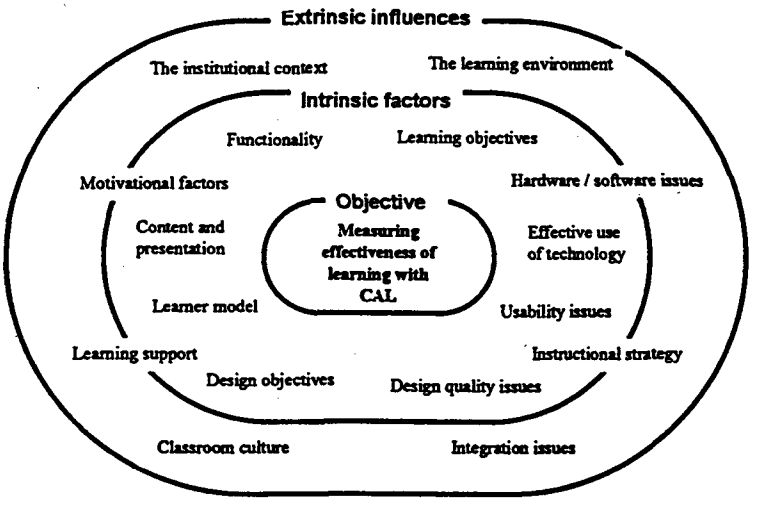

Figure 1: SECAL – a framework for situated evaluation of CAL

SECAL element	Case-specific details
Statement of evaluation objectives	To assess the impact on learning of using CAL to replace some practical laboratory work: • examine the case for investment in more workstations and software licences; • explore the potential for development of stand-alone CAL tutorial exercises; • identify areas for further investigation through qualitative, quantitative or longitudinal studies.
Learning objectives	Students should gain a basis for understanding the biomechanics of movement and the nature and avoidance of sports injuries. The evaluation tasks measured learning by the ability to locate, identify and describe interconnections of various structures.
Hardware/software issues	The product was bought in and so offered no scope for modification. The hardware specification complied with what was already available within the department. However, licence limitations required it to run across a network, and speed, functionality, etc. had not been tested under these conditions.
Effective use of technology	The program presents clear and logically constructed representations of the systems, structures and layers of the human body. Views can be rotated, expanded and dissected in layers. A fair impression of a three-dimensional structure is given by something that is in fact two-dimensional. Access to written and spoken versions of the entire, complex terminology is provided.
Design objectives	These were not measured because the product was bought in. The developers' stated objectives were a useful product for all aspects of medical education.
Design quality, functionality and other usability issues	The program was of accepted high quality, scored well on usability factors, successfully engaged interest, and allowed tasks to be completed in a logical manner.
Instructional strategy	There was no built-in instructional strategy so this was provided by the context of use, i.e. • tailoring the work to an appropriate level of difficulty; • providing 'scaffolding' for novice computer users and those less confident with the subject; • promoting peer support and the benefits of collaborative learning; • using a constructivist approach to learning with CAL.
Learner model	The general learner model was defined by the product developers. It was further specified by contextual factors.
Content and presentation	The content of the program was comprehensive, and presentation was of good, logical standard. Presentation of the program itself was defined by situational factors.
Learning support	Some support was provided by features of the program. Additional sources were from situational factors such as the task, situation within the course, lecture notes, diagrams and the presence of the lecturer.
Motivational factors	• Making learning goals explicitly relevant to the whole course and to assessment requirements; • the attractive appearance of the program; • presenting the task as a series of challenges and providing feedback.
Classroom culture	The group had little previous experience of independent learning, so gradual introduction in a small-group setting was favoured; attention was paid to design of a non-threatening situation for novice users.
Institutional context	There existed no clear institution strategy related to developments in CAL and support was inconsistent

Figure 2: The SECAL framework for evaluating a CAL program in a course on human anatomy

SECAL description

Figure 1 shows the elements included in the SECAL framework, the relative importance of each being determined by the evaluation objectives and the interests being served. It is a simple matter to customize the framework by weighting each element according to its relevance in a particular case.

Integration issues

The elements of the SECAL framework all depend to some extent on CAL being fully integrated into courses. There have been cases where well designed, educationally sound and accessible CAL fails to achieve the success its potential implies. This is often attributable to poor integration strategies, either at institutional or classroom level. Where CAL is simply one available option rather than a compulsory part of a course, take-up rates are frequently poor because there is no compelling reason for students to adopt the new study habits involved. Where staff are not committed to technological advances, little encouragement may be passed on to students. Equally, where institutions do not actively encourage staff to use innovative methods little incentive or support may be available to those who wish to do so. At a broader contextual level, there may be social, political and economic pressures which shape institutional policy on matters of technological change in a positive or negative way. Such influences may seem a long way from the 1990s classroom where students are required to use new technologies as aids to learning and communication, but the measurable effects clearly can extend across this entire range.

Applying SECAL

The scenario in Figure 2 presents an example of the SECAL framework applied to the evaluation of a newly introduced CAL program in a course on human anatomy. The data-collection methods included independent observation, field notes, analysis of task performance, expert opinion and group discussions. A full case study report has been published by Gunn (1996).

Conclusions

The very brief description in Figure 2 of the structure and application of an evaluation framework designed to meet the current requirements of CAL researchers can be summarized in four points:

- generation of a detailed description of the evaluation questions to be answered , i.e. the quality of learning outcomes and the means of effective measurement;
- assessment of the evaluation opportunities presented and methods available in the focus situation;
- consideration of findings in relation to the influence of prevailing situational factors;
- reflection on the evaluation process and study findings with a view to future actions.

CAL is not used alone, and so should not be evaluated in isolation. Attempts to do this have frequently been related to measures of cost-effectiveness or comparative studies. Without minimizing the importance of these issues, they do not address the critical questions about quality of learning and the integration of CAL.

To end on an optimistic note, it was once said of the Model T Ford that if proof had been needed that the motor car provided an economical form of mass transport, it would never have passed the novelty stage. What mattered in the end was that it increased user-choice and provided an enjoyable, effective way to travel, so the economics of production became a priority and eventually made it affordable to the masses. Cost considerations of CAL involve a rather separate and complex set of issues, and no attempt is made to include them in the SECAL framework. The primary focus is on how CAL technology might enhance the quality of learning outcomes in the short term, and in the longer term help to drive major educational and social change.

References

Angus, L. (1986), 'Developments in ethnographic research in education: from interpretative to critical ethnography', *Journal of Research and Development in Education*, 20 (1), 23–31.

Bates, T. (1981), 'Towards a better research framework for evaluating the effectiveness of educational media', *British Journal of Educational Technology*, 12 (3), 215–33.

Brown, J.S., Collins, A. and Duguid, P. (1989), 'Situated cognition and the culture of learning', *Educational Researcher*, 18 (1), 32–42.

Carr, W. and Kemmis, S. (1986), *Becoming Critical: Education, Knowledge and Action Research*, London and Philadelphia: The Falmer Press.

Darby, J. (1992), 'Computers in teaching and learning in UK higher education', *Computers in Education*, 19 (1–2), 1–8.

Dede, C., Fontana L. and White, C. (1993), 'Multimedia, constructivism and higher-order thinking skills' in Maurer, H. (ed.), *Educational Multi Media and Hypermedia*, Boston MA: AACE.

Denzin, N.K. and Lincoln, Y.S. (1994), *Handbook of Qualitative Research*, London: Sage.

Draper, S.W., Brown, M.I., Henderson, F.P. and McAteer, E. (1996), 'Integrative evaluation: an emerging role for classroom studies of CAL', *Computers and Education* (CAL '95 Special Edition).

Ellis, D. (1996), 'Learning with technology (Australia)', Queensland University of Technology via AAHESGIT Electronic Mailing List (LISTPROC@LIST.CREN.NET), archived posting of 1 March 1996.

Elton, L.R.B. and Laurillard, D.M. (1979), 'Trends in research on student learning', *Studies in Higher Education*, 4 (1), 87–102.

Emery, D. (1993), 'Developing effective instructional graphics', *Journal of Interactive Instruction Development*, Fall, 20–4.

Fontana, L., Dede, C., White, C. and Cates, W. (1993), 'Multimedia: a gateway to higher order thinking skills' in Maurer, H. (ed.), *Educational Multi Media and Hypermedia*, Boston MA: AACE.

Geoghegan, W. (1994), 'Whatever happened to instructional technology', paper presented to the 22nd International Conference of the International Business Schools Computing Association, Baltimore, USA. This paper is available at: http://w3.scale.uiuc.edu:80/scale/library/geoghegan/wpi.html.

Glaser, R. and Strauss, A. (1967), *The Discovery of Grounded Theory*, Chicago: Aldine.

Gunn, C. (1995), 'Usability and beyond: evaluating educational effectiveness of computer-based learning' in Gibbs, G. (ed.), *Improving Student Learning Through Assessment and Evaluation*, Oxford: Oxford Centre for Staff Development.

Gunn, C. (1996), *A Framework for Situated Evaluation of Learning in Computer Environments*, Ph.D. thesis, Institute for Computer Based Learning, Heriot-Watt University.

Gunn, C. (1996 in press), 'CAL evaluation: what questions are being answered?', *Computers and Education*, 27 (4).

Gunn, C. and Maxwell, L. (1996), 'CAL in human anatomy', *Journal of Computer Assisted Learning* (in press).

Hammond, N., Gardner, N., Heath, S., Kibby, M., Mayes, T., McAleese, R., Mullings, C. and Trapp, A. (1992), 'Blocks to the effective use of information technology in higher education', *Computers and Education*, 18 (1–3), 155–62.

Jonassen, D.H. (1993), 'The future of hypermedia-based learning environments: problems, prospects and entailments' in Maurer, H. (ed.), *Educational Multi Media and Hypermedia*, Boston MA: AACE.

Keller, J. (1987), 'Strategies for stimulating the motivation to learn', *Performance and Instruction*, 26 (8), 1–7.

Kemmis, S. (1987), 'Schools computing and educational reform' in Bigum, C., Bonser, S., Evans, P., Groundwater-Smith, Grundy, S., Kemmis, S., McKenzie, D., McKinnon, D., O'Connor, M., Straton, R. and Willis, S., *Coming to Terms with Computers in Schools*, report of the Schools Studies of the National Evaluation Study of the Commonwealth Schools Commission's National Computer Education Program, Deakin Institute for Studies in Education, 289–306.

Lybeck , L., Marton F., Stromdahl, H. and Tullberg, A. (1988), 'The phenomenography of the mole concept in chemistry' in Ramsden, P. (ed), *Improving Learning: New Perspectives*, London: Kogan Page.

MacDonald, B. and Jenkins, D. (1979), *Understanding Computer Assisted Learning*, Norwich: University of East Anglia.

Mayer, R.E. and Anderson, R.B. (1991), 'Animations need narrations: an experimental test of a dual-coding hypothesis', *Journal of Educational Psychology*, 83 (4), 484–90.

Paivio, A. (1986), *Mental Representation: A Dual Coding Approach*, Oxford: OUP.

Ramsden, P. (1988), *Improving Learning New Perspectives*, London: Kogan Page.

Sorge, D.H., Russell, J.D. and Weilbaker, G.L. (1994), 'Implementing multimedia based learning: today and tomorrow' in Reisman, S. (ed.), *Multimedia Computing: Preparing for the Twenty-first Century*, Harrisburg and London: Idea Group Publishing.

Spencer, K. (1991), 'Modes, media and methods: the search for educational effectiveness', *British Journal of Educational Technology*, 22 (1), 12–22.

Steinberg, E. (1989), 'Cognition and learner control: a literature review 1977–88', *Journal of Computer Based Instruction*, 16 (4), 117–21.

Yin, R. (1991), *Case Study Research: Design and Methods*, London: Sage.

Zuber-Skerrit, O. (1990), *Action Research for Change and Development*, Brisbane, Australia: CALT, Griffith University.