Phillips, R. (2005) We can't evaluate e-learning if we don't know what we mean by evaluating e-learning.
We can’t evaluate e-learning if we don’t know what we mean by evaluating e-learning!

Rob Phillips
Murdoch University

This issue of Interact is about evaluating the effectiveness of e-learning. Critics of e-learning have regularly noted that there is little evidence of its ability to improve learning outcomes, despite substantial worldwide investment in its development, and its wide uptake. Other articles in this issue will discuss this issue and provide evidence of e-learning ‘working’.

Even when research about e-learning has been published showing that it is effective, or at least no less effective than other approaches, misgivings are held about the validity of that research. E-learning represents a convergence of several fields, including education, computer science, design and media studies. Its multidisciplinary nature and rapid evolution has led to individual researchers taking different approaches, deriving from their individual contexts, to evaluation and research, with little reflection on the appropriateness of their approach.

Research into e-learning is complex, and this has not been sufficiently recognised. Part of the complexity arises from a lack of clarity about the meaning of the terms ‘e-learning’ and ‘evaluation’, and the nature of research into e-learning.

I have recently written about how the term e-learning is used in a one-size-fits-all fashion which confuses discussion about it, and proposed that we should classify e-learning applications in terms of the interactions between: student and student; student and teacher; student and resources; and student and computer (Phillips, 2004). This may help ensure that people are clear about what they’re discussing.

Evaluation

Evaluation is another term that is often used imprecisely, with different meanings in different contexts. For example, evaluation might mean finding out about student perceptions of a university course; it might focus on the suitability of an e-learning product for a teaching purpose, or it might focus on the usability of an e-learning product. In addition, in program evaluation one seeks to provide information to make decisions about an ongoing initiative in public health or education.

In most of these examples, evaluation involves making a judgement about the product or process under consideration, but the product or process needs to be understood well enough to specify the criteria and standards for judgement. While many applications of e-learning are relatively well understood, and criteria can be specified to judge their merit, new developments need research before they can be evaluated. At the same time, the processes used by learners as they use e-learning are context-specific and they therefore need to be investigated by research techniques.

Studies of the effectiveness of e-learning ‘products’ therefore involve a mixture of evaluation (making judgements and decisions about the product) and research (trying to understand how people use the product in order to learn). Because of this, I would argue that it is more appropriate to use the term evaluation research when discussing studies of the effectiveness of e-learning.
The factors discussed above make studies of the effectiveness of e-learning complex and multi-faceted, linking knowledge about e-learning production and knowledge about learning with knowledge about evaluation and research.

**Paradigms of Inquiry**

However, the term *research* is also used imprecisely in academic discourse, because there are many kinds of research. Individual academics work within particular research traditions, with discipline-specific characteristics. For example, a physicist seeks to reproducibly measure a phenomenon in order to support or disprove a theory. A medical scientist might conduct a statistical experiment, with treatment and control groups, because the phenomenon under investigation is complex and is affected by other factors. Results can only be statistically inferred rather than predicted. Further, an anthropologist might seek to explore the behaviour of a community, recognising, and attempting to discount, any preconceptions about that behaviour. Yet again, a computer scientist might develop and implement a new algorithm and test its reliability and efficiency.

These are all different types of research activity, but all share the following characteristics. They seek:

- to satisfy human curiosity and attempt to understand the world we live in;
- to solve problems and manipulate the world we live in.

Research also needs to be disciplined, systematic, explicit and ethical, and “its data, arguments and reasoning be capable of withstanding careful scrutiny by another member of the scientific community” Shulman (1988: 5).

The discipline-specific nature of research leads to, and arises from, different worldviews, or paradigms of inquiry. These are briefly discussed below.

The analytic-empirical-positivist paradigm contains an underlying assumption that people and social interactions can be reduced to quantitatively-defined variables, and that these can be controlled for, or randomised. However, social interactions are too complex to easily control in this way in the real world.

The relativist-constructivist-interpretivist paradigm has a focus on qualitatively describing what is happening in a particular context with little concern for generalisability. Neither paradigm is concerned with change or improvement. Purely descriptive studies may be appropriate when we don’t understand anything about the phenomenon being studied, but this isn’t the case with e-learning.

The critical theory-neomarxist-postmodern paradigm has a focus on changing the ‘world’ not just describing it, and critical reflection is central. The various forms of action inquiry arise from this paradigm.

In studies of e-learning, the focus may be on one or more of: describing and understanding how people use an e-learning product; finding ways to improve an e-learning product; and obtaining evidence about the effectiveness of e-learning. These foci involve elements of the strengths of each of the three paradigms discussed above.

At a recent teaching and learning conference, three participants at a panel session described their studies of student use of e-learning. One study was based on the analytic-empirical-positivist paradigm, one was based on the relativist-constructivist-interpretivist paradigm and one was based on the critical theory-neomarxist-postmodern paradigm. Each study provided useful insights into the e-learning use, but each study also had weaknesses. If multiple approaches had been used, each study would have been more valuable.
A fourth paradigm, the critical realist-pragmatic paradigm combines the most appropriate features of each of the other paradigms in seeking to understand e-learning. In addition, this paradigm enables the researcher to use the full range of available data collection methods, without being tied to methods and methodologies associated with particular inquiry paradigms. Increasingly, this paradigm is seen as the most appropriate to use for e-learning evaluation research.

**Evaluation Research Design**

In conducting evaluation research into e-learning, the most important factor is to reflect on the aims and questions of the research. As Shulman states: “We must first understand our problem, and decide what questions we are asking, then select the mode of disciplined inquiry most appropriate to these questions.” (1988: 15). The focus is therefore on the product or process to be evaluated/ researched, and the questions to ask about it. There is a need for both broad, overarching questions, and associated, specific, answerable questions. In addition, the participants (the people to be investigated/ questioned) and the sources of data should be identified.

The development of both broad and specific evaluation questions is difficult. However, existing knowledge about e-learning can be used to narrow the range of suitable questions. There are two aspects of this: what we know about the design and development of e-learning applications; and what we know about learning in the context of the proposed e-learning use.

The e-learning design and project management literature decomposes the development process into a series of steps, while acknowledging that the process is cyclical and steps may need to be revisited. One characterisation of the development cycle consists of definition of the requirements, specification of the design, production, implementation of the finished product and maintenance. After each phase there is an opportunity to review and revise progress. This is clearly an evaluative process, and it is possible to associate distinct phases of evaluation to each development phase. Figure 1 is a representation of this process proposed by Bain (1999), derived from earlier work by Alexander & Hedberg (1994) and Reeves and Hedberg (2002). Each evaluation phase has a different focus and different questions are asked.

![Figure 1. Mapping of the phases of an e-learning production process to corresponding evaluation phases.](image-url)
The four phases identified by Bain (1999) are:

**Analysis and Design**: analysing the curriculum, analysing teaching and learning activities; and specifying the behaviour of the innovation.

**Development**: finding out if the innovation works in the way it was designed, and what is needed to improve it (closely related to formative evaluation).

**Implementation**: evaluating the effectiveness and viability of the finished product (closely related to summative evaluation).

**Institutionalisation**: evaluating the effects of ongoing use of the innovation within the institution.

Bain (1999) also considered models of learning in developing his framework, distinguishing between the learning environment, the learning process and the learning outcome. The ideas were superimposed on the production phases, resulting in a framework (the learning-centred evaluation framework) which emphasises both aspects of learning and aspects of production. This approach was used in an Australian project aimed at improving e-learning evaluation skills (Phillips, Bain, McNaught, Rice, & Tripp, 2000; Phillips, 2002).

Elements of the framework can be used to develop appropriate specific evaluation questions for the e-learning product under investigation. These can be mapped to the data sources and participants and displayed as an evaluation matrix which summarises the entire plan of the investigation. An example is shown in Table 1. This matrix can be used as a project management tool during the conduct of the study.

Table 1. A representative evaluation matrix illustrating the linkages between the specific questions, the participants and the sources of data.

<table>
<thead>
<tr>
<th>Framework stage</th>
<th>Specific Question</th>
<th>Learner</th>
<th>Teacher</th>
</tr>
</thead>
<tbody>
<tr>
<td>Analysis and design</td>
<td>What are the desired learning outcomes?</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Development</td>
<td>How attractive and accessible do students find the e-learning?</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Development</td>
<td>Does the manner in which students use the e-learning encourage the desired learning process?</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Implementation</td>
<td>What is the evidence that learning is occurring as the e-learning is used?</td>
<td>x</td>
<td>x</td>
</tr>
</tbody>
</table>

**Conclusion**

It is very difficult to describe the complexity of evaluation research into e-learning in a brief newsletter article. I hope I have given you at least an overview of the processes, and an indication of the things to think about. If nothing else, I hope I have given you an appreciation of the special characteristics of research and evaluation into e-learning. I encourage you to carefully reflect on your existing paradigms of inquiry, and their appropriateness, before commencing any e-learning research.
References


