The underpinnings of e-learning evaluation and research

Rob Phillips
Educational Development Unit, Murdoch University, Western Australia, Australia
r.phillips@murdoch.edu.au

Abstract

This paper attempts a fundamental analysis of the nature of research into e-learning. It starts by considering the phenomenon of e-learning, arguing that e-learning is an artificial, designed phenomenon, and that research approaches need to consider how the e-learning environment works and how it can be improved, before we can consider how effective it is. We also establish that e-learning inquiry involves a mixture of evaluation and research, and we discuss this in the context of different disciplinary and interdisciplinary research approaches, arguing that e-learning evaluation research involves a varying mixture of a 'search for fundamental understanding' and 'consideration of use'.

We apply the preceding arguments to the e-learning lifecycle, identifying five different *forms* of evaluation research which are appropriate at various stages: Baseline Analysis, Design Evaluation, Formative Evaluation, Effectiveness Research, and Project-management Evaluation. These forms can be used to guide the design of an e-learning evaluation-research study, in a cyclical research approach. We recognize the strength of design-based research in this context, without claiming that it is appropriate in all circumstances.

The paper then unpacks the process of conducting evaluation research, through the use of *divide-and-conquer* techniques to break down the complexity of an e-learning evaluation-research study. The five forms of evaluation research allow us to conceptualise specific research questions at a particular position in the e-learning lifecycle, and evaluation-research matrices assist us to identify sources of evidence to address these questions.

Introduction

A review of the e-learning literature will reveal that there are diverse approaches to e-learning research across geographical and philosophical boundaries. Individual papers reveal work which is derived from particular paradigmatic and methodological perspectives. A question to ask is whether all of these approaches are valid, and whether they are equally valuable? This paper will address this through a fundamental analysis of the nature of research into e-learning. We start by unpacking the term 'research into e-learning'.

We take the view that research is involved with inquiry into phenomena. It is undertaken primarily to satisfy human curiosity (to come to a more profound understanding of the world) and to solve real-world problems. Academic research is disciplined, systematic, explicit and ethical (Shulman, 1988), distinguished "from other sources of opinion or belief [and] is conducted and reported in such a way that the argument can be painstakingly examined" (Cronbach & Suppes, 1969, p.15). The key point to make here is that research is interested in improved *understanding* of a phenomenon.

We agree with Friesen that e-learning has "come to represent a useful shorthand for a range of different orientations to … the use of technologies in education and learning." (2009, p.4). We choose to use, as one among many, the definition of Littlejohn and Pegler: "the process of learning and teaching with computers and other associated technologies, particularly through use of the internet" (2007, p.15). In this definition, the focus, as it should be, is on learning – learning facilitated through technological tools.

In Phillips, McNaught and Kennedy (2011), we called the software tools which facilitate learning *e-learning artefacts*. We distinguished between three types of e-learning artefacts: interactive learning systems, such as the monolithic interactive multimedia systems of the 1990s; the self-contained learning

objects of the 2000s; and generic learning tools, such as learning management systems. For these tools to assist people to learn, they need to be embedded in an *(e)-learning environment*, which includes specifically-designed learning tasks and interactions.

The phenomenon of e-learning

The next question to ask is what sort of phenomenon is e-learning? A phenomenon is an instance of something – an *artefact* or *event* that is known through the senses. Everyday phenomena, like the seasons, gravity and waves on the ocean *exist naturally* in our world and we can observe them. Other phenomena, such as the inner workings of atoms and molecules, cannot be directly observed, but also exist naturally and their behaviour can be measured in various ways.

However, phenomena are not restricted to things that exist naturally – many phenomena have been created by humans. Simon (1969) distinguished between *natural sciences*, which concern themselves with discovering how existing things work, e.g. physics, biology and anthropology, and *artificial sciences.* Artificial sciences seek to design artefacts, understand and reflect on them, and ultimately improve their design and use, for example in the fields of engineering and architecture. However, designed artefacts do not have to be physical 'things' – they may also be intangible, such as a computer program or an e-learning environment.

Research in the artificial sciences has an extra element not present when researching natural phenomena. With natural phenomena researchers have to take them as they are; but with designed phenomena, there is potential to improve the phenomenon through its design. Thus, research into designed phenomena is not only concerned with the behaviour of that phenomenon, but also with the design and functionality of the artefact which represents the phenomenon.

The implication of this for e-learning is that we need to have a way of knowing that our e-learning artefacts or e-learning environments work properly, and we need to find out how they can be improved. This leads us to the concept of evaluation, particularly formative evaluation.

The terms 'evaluation' and 'research' are often used interchangeably in studies of e-learning. As stated above, research is involved with increasing our understanding of a phenomenon. On the other hand, we see evaluation as gathering information to help make judgements about the value and worth of an object in order to inform decision-making. A distinction is commonly made between two types of evaluation: *formative evaluation*, with a focus on making judgements about improving something; and *summative evaluation*, with a focus on judgements about the merit or worth of something. Evaluation and research can use similar methods to arrive at similar outcomes, but they can be distinguished based on the role of theory in interpreting results and in the way those outcomes are used (Oliver, Harvey, Conole, & Jones, 2007).

In summarising the discussion so far, we can see that e-learning is an artificial phenomenon that results from a design activity, and the outcome of that design activity is an e-learning environment, made up of one or more e-learning artefacts. Given this, any research we conduct needs to consider how the e-learning environment works and how it can be improved, before we can consider how effective it is. This focus on improvement in design activities implies that they are inherently cyclical, and we use the term 'e-learning lifecycle' to refer to the ongoing evolution of a designed e-learning artefact or e-learning environment.

Because an e-learning environment goes through an e-learning lifecycle, we have different goals and ask different questions at different stages of that lifecycle. Some of these questions might be judgment questions, while others might be understanding questions, and we use the term '*evaluation research*' to capture this idea.

Research goals

Shulman (1988) suggested, in his seminal work on research in education, that "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, p.15). In this context, we argue that the nature of the

phenomenon to be studied and the principal outcome sought determine the goals of the research. We recognise, as Reeves (2000a) did, that "the research goals held by any given [educational] researcher are influenced by many factors including the epistemological views of the investigator, his/her research training, and the dominant research paradigms within his/her line of inquiry." (p. 22).

Discussion of paradigmatic and methodological issues is beyond the scope of this paper. Suffice it to say that coherent arguments have been made (Phillips, et al., 2011; Reeves, 2000b) that a pragmatic, eclectic, mixed-methods paradigm of inquiry is the most appropriate approach for evaluation research about the effectiveness of e-learning. By adopting a pragmatic paradigm of inquiry we are able to, at the same time, make judgements about learning environments, seek deeper understanding of them, and aim to improve them. In other words, a pragmatic paradigm of inquiry can enable us to address a range of different research outcomes. The next section discusses the types of research approaches which may be relevant to e-learning.

Research approaches within and across disciplines

The task of clarifying research goals in e-learning studies is complicated by the interdisciplinary and multidisciplinary nature of this emerging field (Friesen, 2009). E-learning is primarily a branch of the discipline of education. However, it also brings in influences from other fields, including computer engineering, information technology, design and media studies. Academic work has often been distinguished across two dimensions: pure versus applied, and hard versus soft (Becher, 1989; Jones, Zenios, & Griffiths, 2004). However, it is possible to identify e-learning research which occupies all quadrants of this space (Phillips, et al., 2011).

Given this, it may be useful to consider other non-disciplinary ways of categorising research, and Stokes' (1997) work on Pasteur's quadrant is helpful in this regard. Stokes questioned the accepted way of classifying research as either pure (discovery) research or applied research, and argued that the distinction was too narrow and did not consider how different types of research might be used. He proposed a two-dimensional model for classifying research (Pasteur's quadrant), where one dimension classified research in terms of the degree to which it reflected a quest for fundamental understanding, and the second dimension focussed on 'consideration of use'.

Earlier, we argued that evaluation research into e-learning involves a mixture of judgements (evaluation) and understanding (research). We consider both of these activities to be inspired by use (in Stokes' terms) because we are studying a learning environment as it is used by learners and through the process of evaluation research we are aiming to better understand how technology can be applied and used. Further, the research component of evaluation research is clearly intent on a quest for fundamental understanding, while the evaluation component is not.

e-learning research goals

Because learning, including e-learning, is complex, the phenomenon under investigation represents an ill-defined problem, and it can be difficult to determine exactly what the research goal is. However, consideration of the state of the phenomenon at different stages of the e-learning lifecycle can help us to identify different general goals:

- design a 'best-practice' e-learning artefact and embed it in a 'best practice' e-learning environment;
- make judgements about how to improve the e-learning environment until it functions as it was designed;
- seek understanding about how students engage with the e-learning environment;
- make judgements about the effectiveness of the e-learning environment; and/or
- seek understanding of how students learn from the e-learning environment.

Some of these are research goals, while others are evaluation goals. A study of the effectiveness of an e-learning environment may quite easily shed light on *how* learners engaged with the designed learning processes to achieve their results, or *why* some learners achieved at different levels, or *how* some learners used the learning environment to achieve a deeper understanding. When it comes to e-learning

investigations, there is typically an ebb and flow between making judgements about the e-learning environment and developing a greater understanding of learning in that environment.

Forms of e-learning evaluation research

The creation of e-learning artefacts and e-learning environments through the e-learning lifecycle arises from a complex, multidisciplinary process, with multiple design, develop, implement and evaluate cycles (England & Finney, 1999; Reeves & Hedberg, 2003; Richey, Klein, & Nelson, 2004). One characterisation of the process is shown in Table 1, which defines seven stages in an idealised e-learning lifecycle, together with the typical activities associated with each stage. These activities follow a lifecycle from analysis, design, development, prototyping to embedding.

Table 1. Components of an idealised e-learning lifecycle.

Cycle	Lifecycle stage	Development activity	Evaluation-research form
0	Analysis of problem	Analyse the learning problem	Baseline analysis
1	Design e-learning artefact	Design the e-learning artefact and document the design	Design evaluation
2	Develop e-learning artefact	Refine the design, develop the e-learning artefact and conduct an initial trial	Project-management evaluation Formative evaluation of the e-learning artefact
3	Design e-learning environment and conduct pilot study	Design and develop an e-learning environment which embeds the e-learning artefact and pilot it	Formative evaluation of the e-learning environment
4	Refine e-learning environment and conduct full trial	Revise the e-learning environment and conduct a full trial with learners	Formative evaluation of the e-learning environment and processes
5	Evaluation research on mature system	Revise the e-learning environment, deploy it to learners and start to understand how it works	Effectiveness research into learning processes and outcomes
6	Evaluation research on mature system	Revise the e-learning environment, deploy it to learners and refine understanding about how it works	Effectiveness research into learning processes and outcomes

Consideration of these activities across the e-learning lifecycle enables us to derive several distinct evaluation-research *forms* (column 4 of Table 1), which are discussed in detail in Phillips et al. (2011).

Baseline analysis

A baseline analysis documents the teaching and learning context. It describes the nature of the teaching and learning problem to be addressed, including a literature review and theoretical positioning, and the pedagogical and technological assumptions of the designers.

Design evaluation

A design evaluation describes and justifies the proposed e-learning environment in terms of curriculum design, learning design, and technology design, and then make judgments about this documented design of the learning environment.

Formative evaluation

Formative evaluation is appropriate at various stages of the e-learning lifecycle, making judgments about, and suggesting improvements to, the e-learning artefact, the e-learning environment and engagement in learning tasks.

Effectiveness research

Once an e-learning environment is functioning as it was designed, effectiveness research becomes appropriate. Effectiveness research mixes components of evaluation and research, seeking to confirm the effectiveness of e-learning environment and develop understanding of how learners engage with learning tasks to demonstrate learning outcomes

An independent, but related, evaluation-research form is **project-management evaluation** Since many e-learning initiatives are developed as projects, it is appropriate to make judgements about, and suggest improvements to, the conduct of that project. Project-management evaluation is primarily interested in processes, rather than outcomes, and it is primarily concerned with formative evaluation, although there are summative elements.

E-learning Research Approaches

The preceding analysis leads us to the conclusion that evaluation research of e-learning is an ongoing cyclical process which is closely related to the cycle of development of an e-learning environment (Bannan-Ritland, 2003; Nieveen, McKenney, & van den Akker, 2006; van den Akker, 1999; Wang & Hannafin, 2005). Moreover, the characteristics of e-learning evaluation research are different at differing stages of the e-learning lifecycle. Systematic, design-based approaches, such as used in engineering (Burkhardt, 2006; Cobb, Confrey, diSessa, Lehrer, & Schauble, 2003; Ross & Morrison, 1989; Salomon, 1991) are more appropriately applied in this area than analytic, scientific approaches. Design-based Reasearch has emerged in recent years as a suitable approach to educational research (van den Akker, Gravemeijer, McKenney, & Nieveen, 2006), in particular e-learning research (Herrington, Reeves, & Oliver, 2009; Reeves, 2006). Design-based research as an approach is consistent with the arguments we have presented earlier about studies of e-learning requiring a mixture of evaluation and research; and with Stokes' views on quest for understanding and consideration of use.

The preceding section has made a strong case for iterative cycles of evaluation research. However, there are many cases where circumstance and/or the research goal dictate a stand-alone approach to evaluation research, which can start at any stage of the e-learning lifecycle. For example, there are many relatively mature learning environments that have been designed in the absence of any research activity. Also, we know that many e-learning projects have relatively short time-frames – either because of grant requirements or the need to have a finished product ready for the next term's teaching.

The cyclical nature of e-learning and the five evaluation research forms can be used as divide-and-conquer techniques to assist in designing an evaluation research study. They prompt the researcher to consider specific research questions for each stage of the e-learning lifecycle. Once research questions have been developed, a second divide-and-conquer technique is the use of evaluation research matrices. These assist a researcher to identify appropriate research methods, and study participants, to provide evidence to

Table 2. Evaluation-research matrix for formative evaluation of an e-learning artefact.

Question \ Data sources	Document Review	Interviews	Surveys	Observation	Think-aloud
• Does the e-learning artefact function as it was designed? What are the bugs?				•	•
• Are interactive learning activities within the e-learning artefact working as designed?		•		•	•
• Can learners use the e-learning artefact easily (can they navigate, gain access to materials, etc.)		•		•	•
• Is the graphic design attractive, approachable and accessible?		•			•

answer each question.

Table 2 is an example evaluation research matrix for the formative evaluation form. It provides a concise mapping of specific research questions and data sources. An evaluation-research matrix can be extended to cover the range of research forms and specific research questions which are appropriate at a particular stage of the e-learning lifecycle. An evaluation research matrix can subsequently be used to develop specific instruments for the study.

Conclusion

In this critical review, we have attempted to shine new light on the problem of researching e-learning by going beyond the numerous implicit assumptions which have informed research practice to date. We have established that e-learning evaluation research is complex, but is essential for designing optimal learning environments and building on the promise of learning technology.

E-learning evaluation research is fundamentally affected by the phenomenon being studied, the goal of the research and the research questions being asked. We argued that e-learning environments are artificial, rather than natural, phenomena, that will evolve, and they need to be researched appropriately. We established that studies of the effectiveness of e-learning environments involve a mixture of evaluation (making judgements) and research (seeking understanding), and an investigation can be placed anywhere along the evaluation–research continuum, depending on its goals. We mapped evaluation research against the e-learning lifecycle, and identified that design-based research is a useful framework for guiding evaluation-research studies.

We identified five evaluation forms that have different emphases when applied at different phases of the e-learning lifecycle. These are: baseline analysis, design evaluation, formative evaluation, effectiveness research, and project-management evaluation. We used these forms to unpack the process of designing evaluation-research studies, and highlighted the use of evaluation-research matrices to break down the complexity of an e-learning evaluation-research study through divide-and-conquer techniques.

In conclusion, rather than applying one-size-fits-all techniques to e-learning research, it may be more appropriate to identify the stage of the e-learning lifecycle that an e-learning environment is in, and conduct research appropriate to that lifecycle stage. Sometimes, this may be to establish that an e-learning environment functions as it was designed; and, at other times, it may be to conduct summative, effectiveness research and a mature e-learning environment.

References

Becher, T. (1989). Academic Tribes and Territories: intellectual enquiry and the cultures of disciplines. Buckingham: Society for Research into Higher Education.

Cronbach, L. J., & Suppes, P. (Eds.). (1969). Research for tomorrow's schools: Disciplined inquiry for education. New York: Macmillan.

England, E., & Finney, A. (1999). Managing Multimedia (2nd ed.). Harlow, England: Addison Wesley Longman.

Friesen, N. (2009). Re-thinking e-learning research: Foundations, methods and practices. New York: Peter Lang.

Jones, C., Zenios, M., & Griffiths, J. (2004). Academic Use of Digital Resources: Disciplinary Differences and the Issue of Progression Proceedings of the 4th Networked Learning Conference. Lancaster University, England, UK.

Littlejohn, A., & Pegler, C. (2007). Preparing for blended e-learning. Abingdon, U.K.: Routledge.

Oliver, M., Harvey, J., Conole, G., & Jones, A. (2007). Evaluation. In G. Conole & M. Oliver (Eds.), Contemporary perspectives in e-learning research: Themes, methods and impact on practice (pp. 203-216). Abingdon, Oxon: Routledge.

Phillips, R. A., McNaught, C., & Kennedy, G. (2011). Evaluating e-learning: Guiding research and practice. New York and London: Routledge.

Reeves, T. C. (2000a). Enhancing the worth of instructional technology research through “design experiments” and other development research strategies. Paper presented at the International Perspectives on Instructional Technology Research for the 21st Century a Symposium sponsored by SIG/Instructional Technology at the Annual Meeting of the American Educational Research Association. April 27, 2000 session 41.29. , New Orleans.

Reeves, T. C. (2000b). Socially Responsible Educational Technology Research. Educational Technology(November-December), 19–28.

Reeves, T. C., & Hedberg, J. G. (2003). Interactive Learning Systems Evaluation. Englewood Cliffs, New Jersey: Educational Technology Publications.

Richey, R. C., Klein, J. D., & Nelson, W. A. (2004). Developmental research: studies of instructional design and development. In D. Jonassen (Ed.), Handbook of research for educational communications and technology (2nd ed., pp. 1099-1130). Mahwah, N.J.: Lawrence Erlbaum.

Shavelson, R. J., Phillips, D. C., Towne, L., & Feuer, M. J. (2003). On the science of education design studies. Educational Researcher, 32(1), 25-28.

Shulman, L. S. (1988). Disciplines of inquiry in education: An overview. In R. M. Jaeger (Ed.), Complementary methods for research in education (pp. 3-17). Washington: AERA.

Simon, H. A. (1969). The sciences of the artificial. Cambridge, MA.: MIT Press.

Stokes, D. E. (1997). Pasteur's Quadrant: basic science and technological innovation. Washington, D.C.: The Brookings Institution.

Acknowledgements

I would like to warmly acknowledge my collaboration with Carmel McNaught and Gregor Kennedy which led to the publication of “Evaluating e-learning: Guiding research and practice”. This paper attempts to summarise some of the key points of that work, and I acknowledge Carmel and Gregor’s contribution to this.