Challenges for Research into Open & Distance Learning:
Doing Things Better – Doing Better Things
Proceedings of the European Distance and E-Learning Network 2014 Research Workshop
Oxford, 27-28 October, 2014
ISBN 978-615-5511-00-4
ISSN: 2707-2819
doi: https://doi.org/10.38069/edenconf-2014-rw-0041

ACHIEVING IMPROVED QUALITY AND VALIDITY: REFRAMING RESEARCH AND EVALUATION OF LEARNING TECHNOLOGIES

Adrian Kirkwood, Linda Price, The Open University, United Kingdom

Introduction

In recent years open and distance education (ODE) has increasingly been equated with digital learning technologies. Through the use of technology, universities in many countries now offer aspects of ODE, whether they are dedicated ODE institutions or teach primarily face-to-face. While the investment in technology has been considerable, findings from research and evaluation studies of learning technologies have had little impact on implementation decisions and teaching practices. Has research contributed to building a body of evidence that can inform and provide a firm foundation for subsequent developments in academic practice? Is evidence being generated and reported that can inform the practices of university teachers and students? Innovation and change should be evidence-informed and we need to ensure that the research and evaluation of learning technology produces findings that can inform other practitioners and policy-makers.

There are concerns about what types of evidence are considered during any implementation decisions (Price & Kirkwood, 2014), misgivings have also been expressed about the lack of a well-established body of evidence and about the quality and validity of many research and evaluation studies. This area of scholarship has been described (Selwyn, 2012) as "notoriously sloppy" and "brimming over with lazily executed 'investigations' and standalone case studies, while also tolerating some highly questionable thinking" (p.213). Many improvements could be made when research and evaluation studies relating to technology and education are conducted.

We have reviewed research literature, reports and case studies relating to learning technology innovations and identified many problems with the ways in which studies were conceived and conducted. Consequently, it is difficult to generalise any findings about effectiveness. We identified issues relating to assumptions and beliefs underpinning research studies and the approaches used to investigate the impact of technologies (Kirkwood & Price, 2013a). Frequently, there was a lack of clarity about the nature of the *enhancement* that technology was intended to bring about and what impact technology would have upon the student learning experience (Kirkwood & Price, 2014). Furthermore, relatively few published accounts of such innovations at university level exhibited a scholarly approach to teaching. Frequently,

interventions appear to be technology-driven rather than being undertaken in response to identified teaching and/or learning concerns (Kirkwood & Price, 2013b).

Here we examine some implications of the shortcomings we identified in published studies. We then suggest ways of avoiding these limitations through taking a more rigorous approach to conceptualising, designing, conducting and reporting research and evaluation studies relating to learning technologies.

How 'fit for purpose' are the research methods utilised?

Research methods are not value-free or neutral: they reflect epistemological positions that determine the scope of inquiries and findings. In other words, there are assumptions and limitations associated with all research methods and approaches and these are often *implicit* or *unstated*. In published research and evaluation studies of the use of technologies for education we have identified:

- A lack of clarity and specificity about what outcomes were expected to be achieved and, therefore, what the focus of the research should have been;
- Narrow or inappropriate conceptions of what constitutes 'scientific' experimentation;
- Poorly conducted 'scientific' experimentation;
- Insufficient attention to the underlying assumptions and models associated with any method of enquiry;
- Unwarranted conclusions being drawn from research findings, often based upon inappropriate expectations.

Before discussing these shortcomings further we explore briefly what we mean by 'rigour' in such research.

What determines 'rigour' in educational research?

We are concerned that much of the published research on learning technologies has been undertaken without a rigorous approach. On the other hand, we are also troubled by the claims made by some researchers that only a highly constrained 'scientific' approach has any validity. A scientific enquiry involves the testing of hypotheses about *why* and/or *how* things happen. It is as much about framing the right questions as it is about adopting any particular approach or methodology. Testing is carried out by carefully collecting evidence that is both appropriate and sufficient to demonstrate whether or not the expected consequences of the hypothesis have happened. If not, the hypothesis must be rejected and a revised hypothesis subjected to scrutiny in a similar manner.

In recent years there has been considerable debate (particularly in the USA) about the extent to which educational research should be more experimental, 'evidence-based' and be directed towards informing policy-makers about 'what works'. Ostensibly, the linking of research and policy-making for practice might seem fairly innocuous. However, it is necessary to examine

ISBN 978-615-5511-00-4

the assumptions and theoretical positions that underlie the various claims in order to understand the nature of the controversy and debate.

Some people claim that generalisable results can only be obtained by the adoption of positivist experimental methods and approaches (Cook, 2002; Slavin, 2002; 2003; Torgerson & Torgerson, 2001). Randomised controlled experimentation, often found in medical research, is proposed as the ideal to be emulated in educational research. It is claimed that research on the use of technology for teaching and learning should involve tightly controlled 'comparative studies' or other forms of experiment. A cumulative synthesis of results from many such studies can be developed through 'systematic reviews' and 'meta analyses' (e.g. Tamim et al., 2011). All require the adoption of a strict experimental approach, the use of quantitative data and statistical analysis techniques. They also relate only to certain types of educational innovation or intervention. Consequently, this narrow and prescriptive view of what constitutes 'scientific' research excludes consideration of any studies that do not meet strict criteria for inclusion. It also reflects just one view of what constitutes education, a highly contested concept.

Many educators and researchers dispute that position for both practical and epistemological reasons (Biesta, 2007; Clegg, 2005; Howe, 2009; Reeves, 2011; Rowbottom & Aiston, 2006; Scriven, 2008; Simons, 2003). We cannot examine those criticisms in detail, but there are many problems to be explored by those aspiring to undertake rigorous experimental research in education. Questions should be asked, such as:

- How similar are the educational and medical contexts – Is it appropriate to equate teaching and learning processes with the treatment of medical conditions?
- How feasible and ethical is it to conduct randomised experiments within education contexts, particularly when (for example at university level) the number of participants tends to be fairly low?
- Exactly what part of the educational process is being investigated when strictly controlled experiments are conducted?

In respect of research on the use of learning technologies there are further contested aspects. For example, the applicability of the much-used 'comparative study' method, which so often leads to 'no significant difference' being the reported outcome. Can that experimental method be an appropriate way to assess innovations aimed at transforming students' learning (rather than maintaining the *status quo* in all respects other than the medium used)? (Kirkwood, 2013) Seeking a suitably rigorous 'scientific' approach, many researchers concentrate their attention on the wrong variables (e.g. instructional delivery modes) rather than on meaningful pedagogical dimensions (Reeves, 2011)[14]. Other research methods and approaches can be suitably rigorous (2011), without invoking narrow experimentation and technological determinism (Oliver, 2011).

ISBN 978-615-5511-00-4

Improving quality and validity

Better conceptualisation of the issues underpinning any study (i.e. the goals, aims and rationale of an innovation; the underlying assumptions about 'teaching', 'learning' and 'enhancement') are essential to improve the quality and validity of research. A better understanding will inform and influence the research approach adopted and the data collection methods involved. It will also clarify what interpretations of the findings are appropriate (or not) at the reporting stage. We suggest the following steps to improve the quality and validity of research.

Ascertain the aims and rationale of the e-Learning project

Why was a technology innovation initiated and implemented? What goals was it trying to achieve? These need to be understood before deciding on the most appropriate research approach and methods. Determine *what* precise form of enhancement is sought from this application of learning technology. For example, is the desired enhancement primarily concerned with issues such as (a) increasing technology use, (b) catering for increased student numbers, (c) improving the circumstances or environment in which educational activities are undertaken, (d) improving teaching practices, or (e) improving – quantitatively and/or qualitatively – student learning outcomes? Researchers must consider *how* any enhancement will be achieved and demonstrated (e.g. greater use, increased time on task, improved student satisfaction with teaching, quantitative and/or qualitative improvements in learning). If the intended enhancement involves 'improvements in learning' how are these conceptualised and how will they be operationalised and demonstrated? These are discussed further in subsequent sections.

Determine the pedagogic purpose of the technology project

A recent critical review of published research and evaluation studies of actual technology interventions (Kirkwood & Price, 2014)[4] found that the primary purpose of each project could be assigned to one of three categories:

- Replicating existing teaching practices;
- Supplementing existing teaching;
- Transforming teaching and/or learning processes and outcomes.

Occasionally the stated outcomes expected of projects were inappropriate for the type of intervention being made. For example, projects that simply replicated existing teaching had unwarranted expectations about the transformation of student *learning*. Simply changing the delivery method does not alter the pedagogic function to any significant extent. A lecture remains a lecture (i.e. a primarily transmissive pedagogic method) whether it is delivered in a lecture-room, as a web-cast to be accessed synchronously and/or asynchronously or as an audio or video podcast accessed 'on demand'.

ISBN 978-615-5511-00-4

Recognise that technologies and tools can be used for multiple educational purposes

Researchers and practitioners must recognise that most technologies/tools (such as blogs, forums, podcasts and wikis) are not associated with just a single 'ideal' role, but can function in a variety of ways for many different educational purposes. The manner in which a technology is used for a particular type of learning activity and anticipated outcomes will reflect the teacher's epistemology and approach to teaching and learning (e.g. transmissive, constructivist, collaborative, etc.). Students' use of a technology in that specific context can differ from that experienced in other contextual circumstances. It is insufficient to describe a technology innovation as being about students 'using a wiki' or 'using a discussion forum'. The educational *purpose* and *mode of deployment* must also be specified and explored.

Determine what benefits are expected to be achieved from a technology intervention and for whom

Try to determine the origins of any learning technology project being investigated. Why was it considered necessary? How was the pre-existing situation to be improved by the use of technology? It is essential to clarify not only the nature of the benefit(s) expected from any project, but also the anticipated beneficiaries. For example, the use of pre-prepared and quality-checked materials and resources can benefit learners, teachers and institutional managers by ensuring that greater consistency and standardisation is achieved. Some other technology-based interventions seek novel outcomes, their primary aim being to enable learners to acquire and develop knowledge and skills that are difficult to achieve by other means. Research and evaluation studies of technology projects should ensure that (a) the full range of relevant benefits and beneficiaries is considered and (b) the methods and approaches used are appropriate. It would be insufficient, for example, for measures of satisfaction to be used to determine whether students' learning had been improved (quantitatively or qualitatively) by a particular intervention.

If some form of learning or teaching enhancement is expected, how is conceptualised in relation to the processes and experiences of those involved?

Is learning enhancement conceived primarily in quantitative terms? For example, many studies make use of the scores or grades achieved by students on specially-devised 'before' and 'after' tests. Others use the normal assessment requirements of a course, usually comparing the results of one 'with technology' cohort of students with another 'without technology' group. Such measures indicate that enhancement is conceived in *quantitative* terms: demonstration of enhancement requires determining whether the technology innovation is associated with more (or less) learning being achieved, through the proxy of test scores.

ISBN 978-615-5511-00-4

Alternatively, an innovation might be seeking to achieve outcomes that are more *qualitative* than quantitative. For example, designing students' use of technology for the purpose of:

- Developing and deepening knowledge and understanding, not simply in terms of knowing more (facts, principle, procedures, etc.), but of knowing differently (more elaborate conceptions, theoretical understanding, etc.);
- Developing an understanding that knowledge is contested (legitimate differing perspectives) rather than absolute;
- Developing a range of 'generic' or 'life' skills, e.g. critical thinking, coping with uncertainty, ability to communicate appropriately with different audiences, working effectively with other people, capacity for reflection upon practice, etc.

Qualitative data collection is almost certainly necessary to demonstrate that the desired qualitative improvement had been brought about.

Whether improvements were conceived in quantitative or qualitative terms, it would never be sufficient to simply ask students whether they felt that their learning had been enhanced. Not only does this not *demonstrate* that any enhancement has been achieved, it also assumes that each student shares their teacher's understanding of what that enhancement actually involves. For example, what valid interpretation can be deduced from aggregating students' responses to the questionnaire item "Do you feel that your learning has been enhanced by x"?

Further, for desired outcomes to be achieved the contextual circumstances must be appropriate. Most notably, the assessment methods and criteria must support those outcomes. The assessment for a course or module constitutes the *de facto* curriculum (Brown, 1997; Havnes, 2004; Rust, 2002; Sambell & McDowell, 1998). Assessment determines what learners do when studying: not only **what** they attend to (and what they ignore), but also **how** they go about learning (Kirkwood & Price, 2008). When students are expected to make use of tools such as wikis, blogs, podcasts, etc. within their normal studies, many will not bother to do so unless using the tool contributes in some way to the course assessment requirements.

Establish what evidence is considered necessary or appropriate to demonstrate the achievement of enhancement(s)?

Any research or evaluation study that aims to gather evidence of better student performance or learning improvement must ensure that relevant forms of data are attained. Kirkpatrick's four-stage evaluation model (Kirkpatrick, 1994) proposes that the effectiveness of education/training is best evaluated at four progressively challenging levels – *Reaction*, *Learning*, *Behaviour* and *Results*. Students' reactions might indicate feelings of satisfaction or positive attitudes, but are never sufficient to determine what learners *know* or what they *can do* as a result of an intervention. 'Learning gains' can only be established by the gathering of appropriate evidence, for example by students demonstrating their understanding or their ability to perform desired tasks or actions. Demonstrating improvements in learning,

ISBN 978-615-5511-00-4

especially those of a qualitative nature, can be difficult and will usually require the use of several data collection methods.

If course assessment is to be used as one form of data collection for a project, it is vital to ensure that the assessment method(s) used is appropriate for the outcomes being sought by the intervention. For example, if a wiki or discussion forum is introduced to encourage students to work collaboratively, the associated course assessment will need to acknowledge and reward group working practices. If assessment remains wholly focused on the outputs of individual students, the 'backwash effect' of assessment (Watkins et al., 2005) will lead learners to revert to competitive rather than collaborative ways of working.

Ensure that the findings justify the conclusions drawn and that no unsubstantiated generalisations or recommendations are made

The findings from a research or evaluation study must substantiate any conclusions or recommendations made. Our literature review (Kirkwood & Price, 2014) found articles in which this was not the case. Favourable reactions from learners (particularly if only responses to multiple-choice questions) should not be presented as the sole source of evidence for learning improvement. In situations where technology has been used to *supplement* existing teaching, any enhanced performance could result from the provision of additional teaching resources or learners spending more time on activities. Similarly, where teaching has been altered significantly to include technology use, researchers must be aware that because changes to several variables have been made, it is inappropriate to claim that just one element (i.e. technology) has been responsible for bringing about any change in outcomes.

Over-generalisation should be of concern. It cannot be assumed that findings from research undertaken in one particular educational context can necessarily be applied in any other context. Often studies provide insufficient details about the context, the design of learning activities, the precise use made of technology, the expected outcomes and the means by which learners were assessed for readers to be able to determine the extent to which findings might be of value elsewhere (Thorpe, 2008).

Maintain an appropriate perspective: clearly differentiate the complexities of the 'here and now' from the idealised 'potential' of any new technology.

All aspects of the educational transaction need to be considered, not just the technology being utilised for teaching and learning. There are two major drawbacks when technology itself is taken as the focus of an investigation. First, there is a tendency to consider the technology as the *agent* of any changes observed, rather than the design of teaching/learning activities make use of technology. The key is how teachers design learning activities appropriate for their students to achieve particular educational outcomes or goals. There are always dangers involved in trying to generalise from one specific context to another. Second, it is always important to consider what innovative role any technology is playing. Is it providing a new means of delivering existing pedagogy (*replicating* or *supplementing* existing teaching), or

ISBN 978-615-5511-00-4

does it contribute to new pedagogical approaches and changes in what and how students learn (*transforming* the learning experience)? Often teachers and researchers are so enthralled by the potential of new technologies that their sense of perspective is impaired. Many investigations fail to take account of and build upon lessons learned from research into the use of educational media and technologies conducted over previous decades, much of which remains highly relevant.

Conclusions

We contend that research and evaluation studies of learning technologies should be conducted with greater rigour and validity. However, it is not a matter of simply following prescriptions about adopting specified research methods or approaches to achieve 'scientific' rigour. It is more about proceeding in a scholarly way, investigating the aims and goals of an intervention in order to pursue **all** relevant aspects of the educational situation and circumstances. Explicit consideration of the assumptions and epistemological models underpinning both the approach to teaching and learning being adopted and the anticipated research methods is essential. The investigation, including any literature review to determine what is already known, should not be focused primarily on the technology being used, but on all relevant aspects of the educational context. All conclusions and recommendations must be supported by evidence and not exaggerated in their claims for applicability in other contexts.

If the guidelines in this presentation are followed, it should contribute to research and evaluation studies achieving higher quality and validity and to results and conclusions that avoid many of the pitfalls and shortcomings that we – and many others – have identified. Consequently, the potential for achieving greater impact will be improved.

References

1. Biesta, G. (2007). Why "what works" won't work: Evidence-based practice and the democratic deficit in educational research. In *Educational Theory, 57(1),* (pp. 1-22).
2. Brown, G. (1997). *Assessing Student Learning in Higher Education.* London, Routledge.
3. Clegg, S. (2005). Evidence-based practice in educational research: A critical realist critique of systematic review. In *British Journal of Sociology of Education, 26,* (pp. 415–428).
4. Cook, T. D. (2002). Randomized experiments in educational policy research: A critical examination of the reasons the educational evaluation community has offered for not doing them. In *Educational Evaluation and Policy Analysis, 24(3),* (pp. 175-199).
5. Havnes, A. (2004). Examination and learning: an activity-theoretical analysis of the relationship between assessment and educational practice. In *Assessment & Evaluation in Higher Education, 29(2),* (pp. 159-176).
6. Howe, K. R. (2009). Positivist dogmas, rhetoric, and the education science question. In *Educational Researcher, 38(6),* (pp. 428-440).

ISBN 978-615-5511-00-4

7. Kirkpatrick, D.L. (1994). *Evaluating training programs.* San Francisco: Berrett-Koehler Publishers.
8. Kirkwood, A. (2013). 'Media are "mere vehicles" – Under-substantiated claims from comparative studies. A review of Richard E. Clark (ed.) Learning from Media: Arguments, Analysis, and Evidence, 2nd Edition'. In *Open Learning, 28(2),* (pp. 153-163).
9. Kirkwood, A. and Price, L. (2008). Assessment and student learning – a fundamental relationship and the role of information and communication technologies. In *Open Learning, 23(1),* (pp. 5-16).
10. Kirkwood, A. and Price, L. (2013a). Examining some assumptions and limitations of research on the effects of emerging technologies for teaching and learning in higher education. In *British Journal of Educational Technology, 44(4),* (pp. 536-543).
11. Kirkwood, A. and Price, L. (2013b). Missing: Evidence of a scholarly approach to teaching and learning with technology in higher education. In *Teaching in Higher Education, 18(3),* (pp. 327-3379.
12. Kirkwood, A. and Price, L. (2014). Technology-enhanced learning and teaching in higher education: What is 'enhanced' and how do we know? A critical literature review. In *Learning, Media and Technology. 39(1),* (pp. 6–36).
13. Oliver, M. (2011). Technological determinism in educational technology research: some alternative ways of thinking about the relationship between learning and technology. In *Journal of Computer Assisted Learning, 27(5),* (pp. 373–384).
14. Price, L. and Kirkwood, A. (2014). Using technology for teaching and learning in higher education: A critical review of the role of evidence in informing practice. In *Higher Education Research & Development, 33(3),* (pp. 549-564).
15. Reeves, T. C. (2011). Can educational research be both rigorous and relevant? In *Educational Designer, 1(4).* Available from: http://www.educationaldesigner.org/ed/volume1/issue4/article13
16. Rowbottom, D.P. and Aiston, S.J. (2006). The myth of 'scientific method' in contemporary educational research. In *Journal of Philosophy of Education, 40(2),* (pp. 137-156).
17. Rust, C. (2002). The impact of assessment on student learning. In *Active Learning in Higher Education, 3(2),* (pp. 145-158).
18. Sambell, K. and McDowell, L. (1998). The construction of the hidden curriculum: messages and meanings in the assessment of student learning. In *Assessment and Evaluation in Higher Education, 23(4),* (pp. 391-402).
19. Scriven, M. (2008). A summative evaluation of RCT methodology: and an alternative approach to causal research. In *Journal of MultiDisciplinary Education, 5(9),* (pp. 11-24).
20. Selwyn, N. (2012). Editorial: Ten suggestions for improving academic research in education and technology. In *Learning, Media and Technology, 37(3),* (pp. 213-219).

ISBN 978-615-5511-00-4

21. Simons, H. (2003). Evidence-based practice: Panacea or over promise? In *Research Papers in Education, 18(4),* (pp. 303-311).
22. Slavin, R. E. (2002). Evidence-based education policies: Transforming educational practices and research. In *Educational Researcher, 31(7),* (pp. 15-21).
23. Slavin, R. E. (2003). A reader's guide to scientifically based research. In *Educational Leadership, 60(5),* (pp. 12–17).
24. Tamim, R.M.; Bernard, R.M.; Borokhovski, E.; Abrami, P.C. and Schmid, R.F. (2011). What forty years of research says about the impact of technology on learning: a second-order meta-analysis and validation study. In *Review of Educational Research, 81(1),* (pp. 4-28).
25. Thorpe, M. (2008). Effective online interaction: Mapping course design to bridge from research to practice. In *Australasian Journal of Educational Technology, 24(1),* (pp. 57-72).
26. Torgerson, C.J. and Torgerson, D.J. (2001). The need for randomised controlled trials in education research. In *British Journal of Educational Studies, 49(3),* (pp. 316-328).
27. Watkins, D.; Dahlin, B. and Ekholm, M. (2005). Awareness of the backwash effect of assessment: A phenomenographic study of the views of Hong Kong and Swedish lecturers. In *Instructional Science, 33(4),* (pp. 283-309).

ISBN 978-615-5511-00-4