Examining some assumptions and limitations of research on the effects of emerging technologies for teaching and learning in higher education

Adrian Kirkwood and Linda Price,
Institute of Educational Technology, The Open University, UK

Abstract

This article examines assumptions and beliefs underpinning research into educational technology. It critically reviews some approaches used to investigate the impact of technologies for teaching and learning. It focuses on comparative studies, performance comparisons and attitudinal studies to illustrate how under-examined assumptions lead to questionable findings. The extent to which it is possible to substantiate some of the claims made about the impact of technologies on the basis of these approaches and methods is questioned. We contend researchers should ensure that they acknowledge underlying assumptions and the limitations imposed by the approach adopted in order to appropriately interpret findings.

Introduction

As new technologies emerge and enter into higher education we must continue to appraise their educational value. However, the way in which we appraise these technologies is important as it influences the results we purport to have found (Oliver, 2011). Researchers’ appraisal methods are underpinned by assumptions about the technology and more significantly about teaching and learning itself (Price & Kirkwood, in press). These assumptions are often underplayed in discourses about the effectiveness of educational technology. They are rarely discussed in articles purporting to have found improvements in learning (Bimber, 1994; Kanuka & Rourke, 2008). This presents variability in interpretation.

Researchers’ beliefs and assumptions shape the research they undertake. Differing epistemological positions reflect how research is conducted. For example, educational researchers may conduct investigations as an objective activity, adopting characteristics of natural or medical sciences. This reflects a positivist epistemology often taking the form of meta-analyses and systematic reviews of quantitative studies (Hattie & Marsh, 1996; Means, Toyama, Murphy, Bakia, & Jones, 2010; Slavin, Lake, Davis, & Madden, 2011; Slavin, 2008; Tamim, Bernard, Borokhovski, Abrami, & Schmid, 2011). Other researchers may adopt a subjective epistemology, seeking understanding. They contend that controlled experimentation is inappropriate for the complex social realities of educational contexts where epistemologies and pedagogies are contested (Clegg, 2005; Elliott, 2001; Hammersley, 1997, 2007; Oakley, 2001). Hence, research methods are not value-free or neutral, but reflect epistemological positions that determine the scope of inquiries and findings.
To illustrate this point, we review some of the methods used to investigate the impact of educational technologies on learning. We question some of the claims made on the basis of the approach adopted and the extent to which these can be substantiated. This critique contributes to current debates about the appraisal of effective educational technologies and their role in enhancing student learning (Oliver, 2011; Oliver et al., 2007).

**Assumptions about learning and teaching**

Interpretations about teaching and learning are frequently taken for granted. However research shows considerable variations in conceptions of teaching (Kember & Kwan, 2000; Prosser, Trigwell, & Taylor, 1994; Samuelowicz & Bain, 1992, 2001). While some teachers have teaching-focused conceptions (i.e. teaching as the transmission of information, skills and attitudes to students), others have learning-focused conceptions (i.e. promoting the development of the students’ own conceptual understanding). Trigwell and Prosser (1996) found that teachers’ conceptions of teaching were commensurate their approaches to teaching. So, teachers with a conception that foregrounds ‘the transmission of knowledge’ are likely to adopt a teacher-centred approach, while those who conceive teaching as ‘promoting conceptual development in learners’ are likely to adopt a learner-centred approach. Teachers’ conceptions of teaching have significant and interrelated impacts upon how they employ technology and upon students’ learning. They also reflects attitudes about agency and whether it is the teacher or the technology that is considered to be significant (Kirkwood & Price, 2012) and this can influence how research is conducted and interpreted, particularly as teachers often conduct research into their own innovations (Hammersley, 1997).

**Comparative studies**

This approach typically involves comparing the outcomes from teaching one group (or more) using some form of technology with those of a control group taught by a more ‘conventional’ method, such as classroom instruction. Apart from the technology, all other aspects of the educational experience are kept identical or as similar as possible. They use the same content, pedagogical approach; they have the same expected learning outcomes and form of assessment. This is in order to establish whether the one factor – the technology – had caused any observed improvements.

This remains a commonly used method in educational technology (Reeves, 2005, 2011; Slavin et al., 2011; Slavin, 2002, 2003, 2008). Means, Toyama, Murphy, Bakia, & Jones (2010) conducted a meta-analysis of 48 studies comparing face-to-face and online or blended learning. In a similar study, Tamim, Bernard, Borokhovski, Abrami and Schmid (2011, p. 5) conducted a meta-analysis of 25 meta-analyses in order to ascertain the impact of technology on student achievement. Neither of these large meta-analyses had any discussion about the comparative research methods paradigm or assumptions that underpinned the design and subsequent interpretation of findings.
The continuing appeal of comparative studies is the apparent simplicity of making a straightforward comparison using a ‘scientific’ method (Oh & Reeves, 2010; Reeves, 2005, 2011). However, this method is not straightforward. ‘True’ experimental comparisons control for a large number of variables and then observe the effects on the dependent variable (Cohen, Manion, & Morrison, 2011, p. 316). This is not easily achievable in real educational contexts as researching human learning is complex (Hammersley, 1997, 2007). More frequently, a quasi-experimental approach is adopted where the teaching received by the experimental group is not just technologically enhanced, but by the very nature of the intervention, supplements or changes the teaching in some manner. This can lead to experimental error as the results are not necessarily due to the manipulation of the independent variable (the technology) alone. This makes causality difficult to establish (Cohen et al., 2011; Joy & Garcia, 2000).

Findings from the majority of comparative studies have resulted in ‘no significant difference’ being found between the effects of the various technologies used for teaching (Arbaugh et al., 2009; Means et al., 2010; Oh & Reeves, 2010; Reeves, 2005, 2011; Russell, 2013). Means et al. (2010) could only find a few studies that met their ‘rigour’ criteria; the other studies could only show ‘modest’ improvements in learning. Reeves (2011, p. 8) observes that comparative studies fail to derive significant results because

most such studies focus on the wrong variables (instructional delivery modes) rather than on meaningful pedagogical dimensions (e.g., alignment of objectives with assessment, pedagogical design factors, time-on-task, learner engagement, and feedback).

Earlier, Schramm (1977, p. 273) observed that

a common report among experimenters is that they find more variance within than between media – meaning that learning seems to be affected more by what is delivered than by the delivery system.

Investigating the impact of technology using the comparative approach, by its very nature, imposes design constraints as the pedagogical components have to remain constant so the effects of the technology can be observed. Hence the technological potential is not advanced or explored. These studies invariably only illustrate findings relating to “doing things better” as opposed to “doing better things” (Reilly, 2005).

In a university context it is usual to consider improvements in learning to be developmental and qualitatively richer. Students are expected not only to develop and deepen their knowledge and understanding, but also to respond constructively to uncertainty, to develop greater self-direction in their learning, and to develop their capacity to participate in a community of practice (Lave and Wenger, 1991). The aspiration here would be to “do better things” (Reilly, 2005). Despite this, many technology enhanced learning studies demonstrate replication of existing teaching practices (Price & Kirkwood, in press). The use of the comparative studies approach
reinforces this finding as they are not suited to investigating the impact of technology on transformational aspects of learning.

Clark (1983) argued that the teaching medium was less significant than the pedagogic or teaching approach when it came to influencing learning. However, he advanced a pervasive analogy based upon the ‘no significant difference’ results frequently found:

The best current evidence is that media are mere vehicles that deliver instruction but do not influence student achievement any more than the truck that delivers our groceries causes changes in our nutrition (1983, p. 445).

The ‘Grocery Truck Analogy’ taken out of its specific context (replication of teaching) could be interpreted as being applicable to all educational situations. However, the evidence had excluded contexts in which technology was used to achieve novel or different learning outcomes. In other words, his generalised assertion – like that of Tamin et al. (2011) – could not be substantiated by the evidence available from comparative studies alone. Clark’s purposeful use of the verb ‘deliver’ indicates that the analogy embodies a transmissive epistemology. Clark’s view of learning concentrates on learners acquiring the knowledge and skills necessary to perform a task through the transmission or delivery of information. This would suggest a conception of learning and teaching with technology that is predicated upon a technologically deterministic perspective, i.e. that the technology in and of itself is the agent of change. This conception is prevalent in assumptions underpinning comparative studies.

Joy and Garcia (2000) argue that the usefulness of comparative studies for predicting learning outcomes is extremely limited due to the need to impose artificial controls to produce results. Constructivist views of learning, aimed at developing student understanding are grounded in very different assumptions and beliefs about the relative roles of instructors, learners and technologies. Such a perspective gives prominence to different research questions that need to be explored through different methodologies (Reeves, 2011).

Performance comparisons

Much educational technology research involves less demanding comparisons between the performance of ‘with technology’ and ‘non-technology’ groups of students (Liao, 1998, 2007; Rosen & Salomon, 2007; Schmid et al., 2009; Sipe & Curlette, 1997; Timmerman & Kruepke, 2006; Torgerson & Elbourne, 2002). Performance is usually compared through normal module assessments or by means of specifically design tests. However, expediency and pragmatism often determines how groups are selected. They might be concurrent groups within the same student cohort, or consecutive cohorts of students taking ostensibly the same module and this can affect the findings given that other factors might affect the results.
When comparing the performance of student groups to determine the effects of any innovation the comparison assumes that the inputs such as resources, learning activities, support, etc. should be equivalent or very similar. If student groups have actually experienced differing amounts of resource input or time spent on tasks, the comparison might provide an indication of improved outcomes, but it cannot be presumed that using technology was responsible for the improvement as the act of changing the resource compromises any claims that can be made about causality.

‘Between group’ performance comparisons tends to assume that student learning gains involve a quantitative improvement, i.e. higher scores achieved reflect more learning (Liao, 1998, 2007). Scouller (1998) has shown that different forms of assessment influences students’ perceptions of the task and their subsequent performance. However, the nature of the assessment itself influences varying student learning outcomes (Laurillard, 1978, 1979, 1984). This suggests that using performance as an assessment of improved student learning has methodological problems. Such methods reveal nothing about whether students achieve longer-lasting gains such as acquiring qualitatively richer or deeper understandings (Dahlgren, 2005; Säljö, 1979) or progressing their intellectual development (Perry, 1970). These kinds of approaches to evaluating student performance similarly reflect what Trigwell et al. (1999) regard as a teacher-focused conception often associated with a transmissive epistemology.

**Self-report questionnaires and attitude scales**

Often researchers try to determine what particular effects innovations have had on learners. For example, how students had used the technology; what types of activity they found most valuable, and what advantages/disadvantages the innovation presented for their study experience, or students attitudes to a particular technological intervention (Cooner, 2010; Copley, 2007; Cramer, Collins, Snider, & Fawcett, 2007; Dalgarno, Bishop, Adlong, & Bedgood Jr, 2009; Elgort, Smith, & Toland, 2008; Evans, 2008; Fernandez, Simo, & Sallan, 2009; Hakkarainen, Saarelainen, & Ruokamo, 2007; Hui, Hu, Clark, Tam, & Milton, 2007; 2008; Sim & Hew, 2010; Sorensen, Twidle, Childs, & Godwin, 2007; Stephenson, Brown, & Griffin, 2008; Tormey & Henchy, 2008; Tynan & Colbran, 2006; Wheeler & Wheeler, 2009; Woo et al., 2008; Wyatt et al., 2010). While such an approach can provide useful information, the outcomes do not of themselves demonstrate that a technological innovation has improved the student learning performance or experience.

Evans (2008, p. 496) conducted a study into the use of podcasts in learning. The questionnaire collected data reflecting students’ experiences and attitudes towards using podcasts. Unfortunately little information was provided regarding any improvements in student’s learning. Cramer et al. (2007, p. 111) conducted a similar study into whether students’ perceived that a Virtual Lecture Hall would enhance their learning. Again, this provided no information about enhancements in learning. While students’ attitudes and opinions are important, other forms of evidence need to be presented in order to conclude whether learning has actually improved. These studies have underlying assumptions in that students’ expressions of attitudes can
be equated with learning ‘enhancement’. This is a dubious interpretation, particularly given that the nature of the enhancement was not specified.

When designing and interpreting the findings from self-report questionnaires, it is easy to assume that all parties share a common understanding. However ‘learning’ and ‘teaching’ are not interpreted in the same way; research has shown considerable variations in interpretation among students and teachers (Kember, 2001; Marton & Säljö, 2005; Trigwell & Prosser, 1996).

The widely used four-level evaluation model proposed by Kirkpatrick (1994) argues that the effectiveness of education or training is best evaluated at four progressively challenging levels: Reaction, Learning, Behaviour and Results. In a critique of the 4-level model, Holton (1996) argues that learner reactions are far less important than the other levels. So while findings suggest that learners value additional flexibility and access of online supplementary resources, research and evaluation studies must go further and investigate any quantitative or qualitative changes in student learning associated with an intervention. Whatever the researcher’s epistemological position or their conception of learning, it is inappropriate to conflate students’ attitudes with their learning development.

**Conclusions**

In this article we critically appraise methods frequently used in educational research. We are not arguing that particular methods are inherently ‘good’ or ‘bad’. Our concern has been to expose the often-implicit assumptions and limitations underpinning methods and to question the extent to which some conclusions are supported by appropriate evidence. Whatever methods researchers employ they should be aware of the underpinning assumptions and limitations of their approach both in relation to the design of the study and in any conclusions that can be drawn from the findings. Interpretations of research need to be cautious as research methods are not epistemologically neutral. Consideration must be given to the extent to which the findings and the design of the study may have been inherently influenced by the research method.

**References**


*et al*