Using concept maps to evaluate student learning and measure e-learning quality

David B Hay and Stylianos Hatzipanagos

King’s Institute of Learning and Teaching (KILT), Kings College London

Abstract:

Measures of student learning are the only authentic indicators of the quality of teaching with technology. The paper will look at a number of frameworks for evaluating e-learning. It will outline some of the implications for using these frameworks to evaluate student learning and will describe the advantages of using concept mapping to measure the quality of e-learning. It will present a study where medical students interacted with e-learning materials designed to teach Magnetic Resonance Imaging (MRI). Their understanding of MRI was measured before and after the course by the use of concept mapping. The results show that many of the student misconceptions were addressed in the course of their learning, but also, that many of the key concepts introduced in the teaching were simply ignored (or sometimes learnt by rote) by the students. These must be acknowledged in the design of electronic teaching materials.

Introduction

The literature on teaching with technology comprises many and varied frameworks for evaluation (e.g., Jones, et al., 1999; Britain & Liber, 2000; Laurillard, 2002). None of these, however, set-out to infer teaching quality from the quality of student learning that they engender. This is an important oversight. Essentially, all evaluation depends on measures of ‘fitness for purpose’ and since teaching has purpose only where it supports learning, learning is the only authentic measure of teaching. That this has not been substantively pursued in teaching evaluation is attributable to two important issues. The first is that learning is commonly deemed too complex a process for empirical measurement (Knight, 2002; Bowden & Marton, 1998). The second issue depends on the obvious truism that while teaching can lead to learning, learning is not a necessary consequence of teaching. Indeed, there is a strand of the literature that suggests that learning is ultimately a consequence of student behaviour, rather than any direct consequence of the teaching they experience (e.g., Tyler, 1949; Shuell, 1986; Biggs, 1996, 2003). This argument should not prohibit attempts to assess teaching through measures of learning.

Learning quality

There is now a small but significant literature on the quality of student learning in higher education (Hay, Kinchin & Lygo-Baker, 2007). That learning can be reduced to measurable criteria (as these authors attest) is attributable to Novak’s work towards a common definition of learning (Novak, 1998). Briefly Novak (1998) states that the quality of change can be located within a continuum from rote to meaningful learning and that meaningful learning can be defined in the following ways:

1. Relevant prior knowledge: the learner must know some information that relates to the new information to be learned in some non-trivial way.
2. Meaningful material: the knowledge to be learned must be relevant to other knowledge and must contain significant concepts and propositions.
3. The learner must choose to learn meaningfully. the learner must consciously and deliberately choose to relate new knowledge to knowledge the learner already knows in some nontrivial way.

This definition is important because it provides a common framework against which learning quality can be measured.

Frameworks for evaluating e-learning

A number of evaluation methodologies have been developed and applied in the context of higher education. Approaches to evaluating e-learning materials vary. A
criterion to differentiate between approaches is whether they are context-based or decontextualised. E.g. Scanlon et al (2000) developed a framework of evaluation (CIAO!) where the notion of “context” is central as social and cultural contexts of learning are a crucial part of that learning. The three components of that framework were (a) context as an indicator of the aims of the e-learning resource and the context of its use, (b) interactions and (c) tutor and student attitudes and outcomes. Within this context, reflective accounts e.g. of how the e-learning resource could be integrated in learning and teaching and experiential accounts of other tutors’ use of the resource help in the evaluation process.

A characteristic of most recent approaches has been an increased emphasis on qualitative rather than quantitative measures. E.g. Squire and Preece (1996) have proposed a framework consisting of eight ‘qualitative’ evaluation questions to consider:

1. Is the complexity of the environment appropriate?
2. Is the learner active?
3. Is fantasy used in an appropriate way?
4. How appropriate is the content to the curriculum?
5. How navigable is the software?
6. What form of learner feedback is provided?
7. What is the level of the learner control?
8. Are learners motivated when they use the software?

Some of these questions could be considered early on at the e-learning selection level and others (e.g. 2, 8) assume interpretive evaluation.

These theoretical frameworks provide scope and context for the evaluation task, however they do not set-out to infer teaching quality from the quality of student learning that they engender. This is an important oversight which concept mapping as an evaluation tool seems to address. Concept maps (Novak 1998) are one strategy that can promote teaching, learning and evaluation in online environments (Daley et al 2007).

The concept mapping method

Concept maps comprise any number of propositional statements each of them made up of paired and linked concepts. Each proposition is a statement of understanding and the validity of each assertion is laid bare. This is a powerful teaching tool since it facilitates the declaration of understanding among teachers and students. Where the same person maps the same topic repeatedly then a comparison of two or more such ‘snap-shots’ facilitates measurement of cognitive change and the quality of change can be assessed against Novak’s original definition of meaningful learning Novak (1998). Those parts of the knowledge structure that are new can be readily differentiated from those that are old and the degree of integration of the new ideas with the extant parts of the prior knowledge structure can be measured.

The study

The data we report here pertain to the evaluation of a single e-learning programme in medical imaging, a series of case studies in medical diagnosis using Magnetic Resonance Imaging (MRI) and a tutorial on the theory of MRI technology. Testing is often part of the e-learning materials development cycle (e.g., Oliver & Herrington, 1995; Georgiadou, Hatzipanagos & Berki, 2005) but our approach was designed to measure the quality of knowledge-change that was a consequence of the students’ e-learning experience. The learning materials that were evaluated belonged to a category described in the literature as narrative systems. Narrative systems support representation in various forms (as opposed to Communication systems that support communication between individuals and groups). An important feature of narrative systems is that they can be used for reception or production of images, sound, text etc. Since learning requires activity, learners should not be on the receiving end for too long without producing some representations of their own, e.g. notes, answers to questions or concept maps.
Methods

Six third year medical students of King’s College London volunteered to take part in the evaluation. Student volunteers were taught to use the concept mapping method (as described by Novak, 1998). Each student was asked to make their own map to describe their prior-knowledge of MRI. Each student then received a CD-ROM of the teaching materials. This comprised instruction of two types. First, it provided a series of interactive diagnostic simulations. Here the students were presented with case studies that required them to ask questions about patient histories and to order diagnostic tests. The programme rewarded their enquiries with medical data and information that led progressively towards a diagnosis. The second part of the package was a tutorial on MRI. Figure 1 is a summary of its structure and content. The tutorial was text-based and annotated with graphics images and diagrams. It was written to teach about MRI functionality and was the ‘core content’ of this part of the e-learning package as a whole.

In all, the simulations and the tutorial were expected to take 6 – 8 hours of study-time and the students were asked to complete the package in the following week. When this time was up, the volunteer group reconvened and the students were reminded briefly of the concept mapping method. Then they were asked to draw their maps of the MRI topic again. This was done without recourse to their first maps or to any other prompts.

Three different approaches were used to score the quality of change in the students’ maps:

1. Analysis of structural change and learning quality: This approach was based on methods previously described by Kinchin, Hay & Adams (2000) and Hay (2007). First, the basic typologies of students’ maps were compared before and after learning to detect gross structural change (Hay & Kinchin, 2006). Second the prior-knowledge maps and the maps made after learning were compared for qualitative measures of change (Hay, 2007).

2. Use of expert terms. This approach comprised a simple analysis of the frequency of concept labels used in the teaching material and in the students’ concept maps after learning. The concept mapping method was used to reduce the text of the e-learning tutorial to a concept map and the concept labels were compared with those used by the students.

3. Expert scores of student maps. An expert teacher and researcher in the field blind marked all of the students’ maps. This was done to measure both the degree of ‘fit’ between student and teacher views, and to give credence to the richness and individuality of student understandings (even where these might not correspond exactly to the content of the teaching). The scoring used five-point Likert scales (from strongly agree to strongly disagree) and measured the following constructs:
   - Conceptual richness (the richness and relevance of ideas used to describe the topic).
   - Linkage and link quality (the richness and appropriateness of concept linkage).
   - Evidence of understanding (the degree to which the propositions describe understanding).
   - Hierarchy and structure (the complexity and validity of the structural representation of knowledge).

All twelve of the student’s maps (six prior knowledge maps and six maps after learning) were given to the expert in random order and all of the statements were ordered at random in each individual score sheet. The expert was also asked to identify any significant mis-conceptions in the student’s maps and to list any concepts or propositions that were conspicuous by their absence.

Results and discussion

The student learning measures obtained using each of the methods showed that the
students' knowledge structures were richer and more explanatory after e-learning. Nevertheless, some students had clearly learnt more than others. Three satisfied the criteria for meaningful learning and three did not (method one). Among those that did not, two showed rote learning and one showed rote learning outcomes mixed with non-learning.

The results obtained from method two offered less substantial evidence for the quality of the students’ e-learning. The teaching materials comprised 78 different concepts. However, only 28 of these were represented in the student’s post learning maps. Furthermore, all of these were derived from just three of the six parts of the tutorial. Not one term used in some parts 2, 3 or 5 of the tutorial found a way into any of the students’ maps. Moreover, learning directly attributable to parts 4 and 6 of the tutorial was confined to only three students. The other three used ‘taught’ concepts from part 1 alone. Even the two students who used more newly taught ideas than anyone else, did so through surface-level addition.

Previously, Hay, Wells & Kinchin (2007) have suggested that prior knowledge structure is the best predictor of the meaning-making that can occur in the course of learning. Kinchin, Hay and Adams (2000) suggest that cognitive restructure is a key component of meaningful learning and Hay and Kinchin (2006) have shown how spoke structures in particular are amenable to change. While we suggest here that the quality of students prior knowledge is likely to determine the meaning-making that they can make from the teaching they experience, it is also likely that the different ways that students go about organising their knowledge can be more or less amenable to learning.

**An authentic pedagogy of e-learning**

Concept mapping has a unique role to play in the pedagogy of higher education. On the one hand it provides an epistemological framework for the practice of teaching and on the other, it is grounded in a theory of learning (e.g. Ausubel, 1963, 1968; Ausubel et al, 1978). Our results lead us to suggest measures of student prior-knowledge should be a key step in e-learning materials design. It is only out of prior knowledge that students are able to make meaning of the new topics that they are taught. If content is ‘locked up’ in e-learning material in ways that are not accessible to students then non-learning or rote-learning outcomes must be expected. Our data also suggest that where learning is meaningful it is a consequence of individual thinking, development and synthesis. It is not a clear and attributable outcome of teaching. This means that good teaching will always be difficult to measure directly, but it also has significant implications for teaching practice. If meaning-making is dependent on the behaviours of learners then teachers should do at least as much to encourage and support student-centred learning activities as they do to develop and deliver content. Good teaching is likely to comprise measures of student prior-knowledge and repeated measurement of newly emerging understanding.

**Conclusion**

The paper discussed a framework to support evaluation of learning materials. Approaches so far are mostly in ‘sound, pedagogically tested theories’ and robust theoretical evaluation frameworks. None of these, however, set out to infer teaching quality from the quality of student learning that they engender.

This paper supports the view that student-learning quality is a product of student activities and behaviour (rather than any direct consequence of taught content). Nevertheless, we also suggest that the ways that teachers present new knowledge to students can have an important bearing on the ways that their students learn. Our data suggest that measurement of student prior-knowledge and subsequent teaching design is key to effective e-learning materials design.

**References:**


Here parallel imaging was described.

This part of the text described the origins of the MRI technique and its use in imaging of the great blood vessels.

Technical issues such as flip angle, phase contrast, inversion time, and temporal resolution were discussed here.

Figure 1: A summary of content in the e-learning material

Figure 2: One student’s concept maps before (left) and after (right) e-learning

Significant misconceptions are shown in the shaded zones.