

Evaluating a virtual learning environment in the context of its community of practice

Rachel Ellaway^{*}, David Dewhurst & Hamish McLeod
University of Edinburgh, UK

The evaluation of virtual learning environments (VLEs) and similar applications has, to date, largely consisted of checklists of system features, phenomenological studies or measures of specific forms of educational efficacy. Although these approaches offer some value, they are unable to capture the complex and holistic nature of a group of individuals using a common system to support the wide range of activities that make up a course or programme of study over time. This paper employs Wenger's theories of 'communities of practice' to provide a formal structure for looking at how a VLE supports a pre-existing course community. Wenger proposes a Learning Architecture Framework for a learning community of practice, which the authors have taken to provide an evaluation framework. This approach is complementary to both the holistic and complex natures of course environments, in that particular VLE affordances are less important than the activities of the course community in respect of the system. Thus, the VLE's efficacy in its context of use is the prime area of investigation rather than a reductionist analysis of its tools and components. An example of this approach in use is presented, evaluating the VLE that supports the undergraduate medical course at the University of Edinburgh. The paper provides a theoretical grounding, derives an evaluation instrument, analyses the efficacy and validity of the instrument in practice and draws conclusions as to how and where it may best be used.

Introduction

Virtual learning environments (VLEs), and systems like them, provide 'the 'online' interactions of various kinds which can take place between learners and tutors, including online learning' (JISC and UCISA, 2003). These systems all share a common thread; they can take on many roles and they can support a wide range of

^{*}Corresponding author. Learning Technology Section, College of Medicine and Veterinary Medicine, University of Edinburgh, 15 George Square, Edinburgh EH8 9XD, UK. Email: Rachel.Ellaway@ed.ac.uk

educational, administrative and logistical processes, each of which is able to interact and integrate with the others.

Because of this integration, because VLEs can be used in many different ways, and because much that was implicit in the traditional learning environment becomes explicit in its online equivalent, the evaluation of VLEs has proved to be a particularly complex problem. Furthermore, because of the sheer scale, complexity and cost of VLEs, their adoption and use is increasingly undertaken at an institutional level and any subsequent evaluation, if it is not done at the level of the individual learner, is most often also undertaken at this institutional level. Between the micro and macro approaches are levels that remain relatively disregarded, those of the programme of study or cognate discipline area, which can, in some cases, be usefully modelled as a distinct 'community of practice'.

This paper proposes a holistic approach to evaluating VLEs in the context of the community of practice in which they are used, where such a community already exists. This is based around the 'learning architecture framework' (LAF) proposed by Etienne Wenger (1998). This approach is practitioner oriented and was developed in response to requirements for evaluating programme-wide VLEs in integrated subjects such as medicine and veterinary medicine.

Evaluating VLEs

Although there has been much published on evaluative work on VLEs, this has until recently rarely gone further than analysing their various features and functions (see CHEST MLE/VLE comparison grid;¹ EDUTOOLS comparison grid;² or Jenkins *et al.*, 2001) or the phenomenology of VLEs in use (Lee & Thompson, 1999; Barajas & Owen, 2000; Richardson & Turner, 2000). Occasionally more sophisticated approaches to evaluating VLEs have emerged (Britain & Liber, 1999; Koper, 2000) which take a more grounded and pedagogically orientated approach, but which continue to orientate towards predictive and intrinsic properties of the VLE. In presuming that a VLE has intrinsic properties, that the context into which a VLE will be deployed is neutral and that any given VLE will automatically deliver predictable benefits (or otherwise) into that context, the predictive approach is significantly limited in providing a useful perspective of a VLE in a grounded course context. It is important to note that most of these approaches have been directed towards a novice audience looking for the best evidence or advice available to help them select a suitable system to meet their needs.

A different approach to prospective and predictive models is the evaluation of a 'VLE-in-use', which investigates the unique properties and dynamics of a course-VLE instance. This is a variation on 'situated' or 'holistic' evaluation (Bruce *et al.*, 1993). Such approaches are now beginning to appear in the literature, for instance using grounded theory (Alsop & Tompsett, 2002).

In introducing a typology of approaches to evaluating learning technologies Oliver (1998) identifies 'holistic evaluations' as starting from the position that introducing technology to educational settings will tend to alter learning outcomes rather than just

the quantity or quality of what is learnt. The holistic approach seeks to encompass broader aspects of learning technology such as its social or organizational dimensions as well the immediate focus of interest. Oliver describes this methodology as seeking:

to identify the positive and negative aspects of technology use, as compared with traditional teaching methods, to build categories from these data, and then to statistically analyse the response patterns in order to arrive at generalizable findings.

A situated and/or holistic 'VLE-in-use' approach is therefore likely to be a more appropriate methodology to adopt to unlock the nature and value of our VLEs in their complex and multifactorial relationships with their course contexts.

Asking the right questions

Whether used for distance learning or deployed in support of on-campus courses, and whether a commercial 'off-the-shelf' or a 'home-grown' system is being used, VLEs share two essential characteristics; multiple systems integration and a course or module focus.

It might be expected that any given VLE might be employed in different ways within different course contexts. In the same way, a particular course could be expected to use different VLEs in different ways. Thus, the pairing of a course and a VLE should be understood as a unique instance, with a unique set of characteristics, rather than as two separate entities. The effectiveness and value of a VLE to a course is therefore not an inherent property of the VLE software but depends on its use in facilitating and mediating the needs and activities of a particular course. Extending this argument further it might be concluded that all VLE functions exist in a 'blended' relationship with human activities, independent of whether they are the primary delivery medium (e.g. distance courses) or one among many (e.g. on-campus courses).

This is of course true of any technology, but it is of key importance when considering complex situations where the permutations of how different kinds of affordances can be taken up far exceed the prime designed affordances of any given technology. Added to this is a need to consider the adaptive nature of human practices to the affordances of the available technologies. If a thing can be done, not only will the direct affordances it offers tend to change our practices, but, in doing so, it will most likely change what we want to do, and what value we attach to these different activities and functions (Graham, 1999).

The defining aspect is therefore how such complex technologies are used in specific circumstances. The question which should be asked about a VLE is not 'what can it do?' but rather 'what is it doing?', thereby focusing on its function and role in the situated educational context.

It is often the case, however, even for the best managed courses, that over time their ethos and procedures become blurred, sometimes to the point where there is no clear or common understanding of what the course is about and, more importantly, exactly how it works. Furthermore the deployment of a VLE into any course context initiates

a complex set of exchanges, each component shaping the other in a cycle of mutual adaptation.

the individual organization is neither merely a passive receiver of predetermined technological artefacts nor an autonomous controller of technological change. Rather, in organizing the flows of knowledge and resources within and between groups, organizations shape the technology process at the same time as it shapes them. (Scarborough & Corbett, 1992, p. 10)

There may therefore be no single reliable or meaningful way to independently measure a course's dynamics or the relevant contextualized affordances of a VLE. However, given a common set of defining criteria or factors, a VLE may be evaluated within the specific course context in which it is used, avoiding the problems of independent analysis and benefiting from the situatedness of such an approach. Because of the complex nature of this VLE problem space it is unlikely that any one technique will be effective in evaluating a VLE fully. However, a triangulation approach (Breen *et al.*, 1998), employing a range of techniques and perspectives is likely to provide a better solution. The rest of this paper describes the creation and use of an evaluation instrument based on Wenger's theories on 'communities of practice' that can provide a strong contributing dimension to this kind of analysis.

Communities of practice and VLEs

Originally focusing on ideas of apprenticeship, the development of theories of 'communities of practice' coalesced around the concepts of legitimate peripheral participation in a community of practice identified by Lave and Wenger (1991). Legitimate peripheral participation in a community of practice is centred on the notion that:

learners inevitably participate in communities of practitioners and that mastery of knowledge and skill requires newcomers to move toward full participation in the sociocultural practices of a community. (Lave & Wenger, 1991, p. 29)

In 'Communities of Practice', Wenger (1998) proceeded to argue that current orthodox approaches to learning that are based on concepts of individual learners learning in prescribed ways causally linked to teaching are redundant. Wenger proposed instead a 'social theory of learning', which is based upon learning as individual engagement and participation in a community of practice.

Wenger's theories can be particularly relevant in modelling learning environments where they encompass a pre-existing learning community of students, teachers/tutors, support staff and potentially many other roles and groups. Furthermore, any participant may adopt or change roles; students may be involved in teaching each other, teachers may become learners, support and administration responsibilities may fall to different participants at different times and so on. All of this activity is in turn informed by the socio-cultural norms and values inherent in the practice and the related social contexts in which it is situated. If this is the case then such a course may be modelled as a community of practice, and indeed, its component parts (such as

modules of study or groupings such as ‘students’) may themselves constitute subsidiary communities of practice. The relevance of this model depends on the degree and coherence of shared purpose, meaning, activity and identity across the course community. For instance in a modular arts or science context the model may be expected to be weak as students pursue individual patterns of cross-disciplinary study while in an integrated vocational context such as law or teaching the model would be expected to be more relevant.

There are existing studies that have applied Wenger’s theories in the context of VLEs (e.g. Rogers, 2000; Chalk, 2001). These have, however, tended to take Wenger’s general topics of ‘mutual engagement’, ‘joint enterprise’ and ‘shared repertoire’ as the basis for their work rather than anything more structured. Their focus has also tended to fall uneasily between the learner and the environment without clearly identifying one from the other.

However, following his discussions of the general dynamics and characteristics of communities of practice, Wenger goes on to formalize these dynamics in his ‘Learning Architecture Framework’ (LAF) for a learning community of practice. This framework has the following defining characteristics or properties (Wenger, 1998, pp. 237–239):

Facilities of engagement

- Mutuality—interactions, joint tasks, help, encounters across boundaries, degrees of belonging.
- Competence—opportunities to develop and test competences, devise solutions, make decisions.
- Continuity—repositories, documentation, tracking, ‘participative memory’, storytelling, ‘paradigmatic trajectories’.

Facilities of imagination

- Orientation—location in space, time, meaning and power.
- Reflection—models and patterns, opportunities for engaging with other practices or break rhythm with the community mainstream.
- Exploration—trying things out, simulations, play.

Facilities of alignment

- Convergence—common focus or cause, direction, vision, values, principles.
- Coordination—procedures, plans, schedules, deadlines, communication channels, boundary encounters and brokers.
- Jurisdiction—policies, contracts, rules, authority, arbitration, mediation.

By accepting these as key properties of a learning community of practice, this framework can be used as the basis of a more structured evaluation methodology,

evaluating a VLE in its context of use. It is proposed that by using this framework a VLE can be evaluated in terms of its success and value in supporting these nine properties in the context of the community of practice that employs it.

A note of caution should be added at this point. The approach advocated in this paper is a descriptive post-hoc evaluation model and is not intended as a template for designing a VLE. As Schwen and Hara observe:

while Wenger's work is a provocative ideal to achieve and useful as a dialogue between the designers and client systems, it is not a recipe for construction of such phenomena (Schwen & Hara, 2003, p. 262)

This approach is also intended for use where a VLE is one medium for course delivery amongst many. Some approaches that draw upon the principles of communities of practice are predicated on the community of practice being either fully or predominantly online (Notess & Plaskoff, 2004). It is more usual in higher education for a VLE to provide scaffolding and support within a multi-modal environment (Ellaway *et al.*, 2003).

Method: developing a VLE evaluation tool based on the LAF

The first stage of the development process was, starting with Wenger's LAF, to move from the general component factors of the LAF to increasingly specific questions aimed at evaluating VLEs. The first three steps of this process are shown in Table 1 and were at this stage fully derived from Wenger.

The second stage was to extend Wenger's theories to develop a pool of VLE-oriented questions based on the 'specific aspects' column of Table 1. This pool of questions was piloted with a variety of members of the target learning community and, as a result, a number of questions were combined, rephrased or omitted. A particular outcome of this piloting was the development of a three-stem structure for each question, based on general effectiveness, personal utility and personal value. The questions were then rephrased as value statements and participant response option was structured against Likert scales. Rather than creating new scales, the Likert scales were selected from those available in the online evaluation system that was to be used in delivering the instrument. The instrument was piloted again and further refinements and adjustments made.

In addition to questions derived from the LAF, a number of questions were included to verify responses against other data (such as server logs) and to act as 'consistency traps'. When completed the evaluation instrument comprised of 60 items (see Table 2). The instrument was designed to be administered to all of the VLE's users, irrespective of their role in the learning community supported by the VLE; whether student, academic or support staff.

This is, as was noted earlier, an inherently 'situated' and faceted approach; the nine-point LAF has no inherent hierarchy or ranking of importance or relevance. These can only be judged or evaluated in the context at hand. Indeed, any given course context may well contain distinct constituent communities of practice (such as

Table 1. First steps in developing Wenger's Learning Architecture into an evaluation instrument (after Wenger, 1998, pp. 237–239)

General factors	General questions	Specific aspects
Mutuality	Does/will the system support and facilitate the required mutuality and how important is this?	Interaction Joint tasks Peripherality
Competence	Does/will the system support and facilitate the required competences and how important is this?	Initiative and knowledgeability Accountability Tools
Continuity	Does/will the system support and facilitate the required continuity and how important is this?	Reificative memory Participative memory
Orientation	Does/will the system support and facilitate the required orientation and how important is this?	Location in space Location in time Location in meaning Location in power
Reflection	Does/will the system support and facilitate the required reflection and how important is this?	Models and representations Comparisons Time off
Exploration	Does/will the system support and facilitate the required exploration and how important is this?	Scenarios Simulations Practicum
Convergence	Does/will the system support and facilitate the required convergence and how important is this?	Focus, vision and values Leadership
Coordination	Does/will the system support and facilitate the required coordination and how important is this?	Standards and methods Communication Boundary facilities Feedback facilities
Jurisdiction	Does/will the system support and facilitate the required jurisdiction and how important is this?	Policies, mediation, arbitration and authority

staff or students) that themselves hold contrasting and conflicting perspectives and value systems within the broader course community of practice.

In taking this approach the authors are making the following assumptions:

- that Wenger's theories adequately model a community of practice;
- that the subject area or discipline has a strong identity as a community of practice;
- that the community of practice encompasses the whole course.

This approach does not seek to test theories of communities of practice. Rather it assumes a pre-existing course community of practice as a given reference point and thereby evaluates a VLE by its ability to support that course community of practice.

Table 2. Learning Design Questionnaire—generic format showing response type and mapping to the Learning Architecture Framework

No.	Root	Type	LAF Mapping
1	How effective is the system's engagement with the course in general?	QA	General
2	How useful is the system at engaging you with the course?	QA	
3	The system's support of my engagement with the course is important to me ...	SA	
4	How effective is the system in general at supporting interactions with students and staff on the course?	QA	Mutuality, competence, continuity, coordination
5	How useful is the system in supporting your interactions with students and staff on the course?	QA	
6	The system's support of my interactions with students and staff is important to me ...	SA	
7	How effective in general is the system at supporting collaborative activities required by the course?	QA	Mutuality, competence, continuity, exploration
8	How effective is the system at supporting the collaborative activities you are involved in?	QA	
9	The system's support of the courses' collaborative activities is important to me ...	SA	
10	How effective is the system in general at providing the course information and help required for the course?	QA	Competence, continuity, orientation, convergence, coordination
11	How useful is the system at providing the course information and help that you require to participate fully in the course?	QA	
12	The system's provision of course information and help is important to me ...	SA	
13	How useful is the system in general at interacting with University services and systems beyond the course?	QA	Mutuality, reflection, coordination
14	How useful is the system at supporting your interactions with University services and systems beyond the course?	QA	
15	The system's support of my interactions with University services and systems beyond the course is important to me ...	SA	
16	How effective is the system in general at supporting assessment in the course?	QA	Competence, continuity, orientation, coordination, jurisdiction
17	How useful is the system at supporting your assessment needs in the course?	QA	
18	The system's support of assessment is important to me ...	SA	
19	How effective is the system at providing the courses' guidelines, rules and regulations?	QA	Competence, continuity, convergence, coordination, jurisdiction
20	How useful is The system's provision of guidelines, rules and regulations you require?	QA	
21	The provision of guidelines, rules and regulations by the system is important to me ...	SA	
22	How effective are the tools provided by the system?	QA	Mutuality, competence, continuity
23	How useful to you are the tools provided by The system?	QA	
24	The provision of tools by the system is important to me ...	SA	

Table 2. Continued

No.	Root	Type	LAF Mapping
25	How effectively in general does the system support progression through the course?	QA	
26	How useful is the system's support of your progression through the course?	QA	Continuity, orientation, convergence
27	The system's support of my progression through the course is important to me ...	SA	
28	How effective in general is the system at supporting out of hours working?	QA	
29	How useful is the system at supporting your need to work out of hours?	QA	Continuity, orientation, reflection
30	The system's support of my work out of hours is important to me ...	SA	
31	How effective is the system in general at supporting teaching and learning activities at different locations?	QA	
32	How useful is the system at supporting your teaching and learning activities at different locations?	QA	Continuity, orientation, reflection
33	The system's support of my teaching and learning activities at different locations is important to me ...	SA	
34	How effective is the system at providing timetabling and scheduling information?	QA	
35	How useful is the system at providing the timetabling and scheduling information you require?	QA	Continuity, orientation, coordination
36	The system's provision of timetabling and scheduling information is important to me ...	SA	
37	How effective is the system in general at providing secondary learning materials?	QA	
38	How effective is the system at providing you with secondary learning materials?	QA	Competence, orientation, exploration
39	The system's provision of secondary learning materials is important to me ...	SA	
40	How effective is the system at providing access to materials and resources that help with the reflective aspects of the course?	QA	
41	How useful is the system at providing materials and resources that help you with reflective aspects of the course?	QA	Competence, orientation, reflection, exploration
42	The system's provision of access to materials and resources that help with reflective aspects of the course is important to me ...	SA	
43	To what degree does the system embody the focus, vision and values inherent in the course?	QA	
44	How useful is the system's embodiment of the focus, vision and values inherent in the?	QA	Orientation, convergence, coordination, jurisdiction
45	The system's embodiment of the focus, vision and values inherent in the course is important to me ...	SA	
46	How effective is the system at supporting the educational practices and methods of the course?	QA	
47	How useful is the system at supporting the educational practices and methods of the course?	QA	Mutuality, competence, convergence, coordination
48	The system's support of the educational practices and methods of the course is important to me ...	SA	

Table 2. Continued

No.	Root	Type	LAF Mapping
49	How effective is the system at supporting feedback and evaluation within the course?	QA	
50	How useful to you is the system at supporting feedback and evaluation within the course?	QA	Mutuality, continuity, coordination
51	The system's support of feedback and evaluation within the course is important to me ...	SA	
52	How effective is the system at tracking student and staff use of The system?	QA	
53	How useful to you is the system's ability to track student and staff use of The system?	QA	Competence, continuity, jurisdiction
54	The ability to track student and staff use of the system is important to me ...	SA	
55	Overall I think the system is a very useful system in helping me engage with the course ...	SA	
56	Overall I think the system is a very valuable system in helping me engage with the course ...	SA	General
57	Overall I think the system is a reliable system in helping me engage with the course ...	SA	
58	How often do you use the system?	ON	
59	How responsive to requests for help and/or support is the system?	QA	
60	Are there any aspects of the system that you think should be added to, improved or changed to make the system more useful to you?	FT	

Response types: QA = excellent to awful; SA = strongly agree to strongly disagree; ON = all the time to never; FT = free text; YN = yes/no.

Using the LAF evaluation instrument

The LAF evaluation instrument was used to evaluate the 'Edinburgh Electronic Medical Curriculum' (EEMeC),³ a purpose-built VLE system supporting the undergraduate medical course at the University of Edinburgh. It was considered that this course had a strong existing course community of practice: students followed a common and integrated programme of study that was not shared with any other students; the programme of study was intrinsically focused on inducting students into medical practice; it was taught by practicing clinicians (often in real clinical contexts), it had a strong socializing agenda, and it had a continuous history going back over more than two centuries..

EEMeC was already well established across the course supporting approximately 1,200 students and approximately 700 staff across all five years of the programme. The development and characteristics of EEMeC have been described elsewhere (Warren *et al.*, 2002; Ellaway *et al.*, 2003).

The survey instrument was deployed using EEMeC's own 'evaluation engine' which allows staff to create, schedule, deliver, record and analyse questionnaires online (Wylde *et al.*, 2003). This is done by mapping different copies of the instrument in the VLE database to the groups that would receive them. By setting a start and end time, when a user logged in to EEMeC with the period set, the system would check to see that an uncompleted questionnaire was set for the user's group and, if this was the case, would present the questionnaire to them in a pop-up window. Although a log of who had completed questionnaires was kept, this was separated from the responses so that they were anonymized at the point of storage and only one response per individual was permitted.

The period set for delivery was from 10 April to 30 April 2003, a period that mapped on to different years based on their term or rotation schedule:

- Year 1—on vacation to 15th April then starting term 3.
- Year 2—on vacation to 15th April then starting term 3.
- Year 3—new clinical rotation started 10th April.
- Year 4—in middle of clinical rotation (24/2 to 30/5).
- Year 5—in middle of clinical rotation (31/3 to 23/5).

It is important to note that this was the first time that staff had been surveyed in this fashion.

The response rates (shown in detail in Table 3) were high overall although the staff responses were particularly low. The figure of 699 staff includes 50 or so guest logins and a large number of clinical and related staff who have a relatively peripheral engagement with the course. It is a peculiarity of medical education that a large number of clinical staff will be involved in teaching but only for a fraction of the working year. Thus, despite the high potential numbers of staff in the course, at any given time only a relatively low number are actively engaged in teaching. This explains to a major degree the relatively poor responses in the staff cohort.

There is also a notable lower response rate in year 3 with only just over half of the year responding relative to that in year 4. In the 2002–3 academic session, EEMeC was progressively less important to the students' engagement with the course in later years, as there were less tools and materials provided, and it was expected that there

Table 3. Responses for all cohorts

Course role	Population	Returns	% returns
Year 1 students	236	207	87
Year 2 students	214	186	87
Year 3 students	258	142	55
Year 4 students	221	192	87
Year 5 students	176	120	68
All staff	699	45	6

would be a gradual decrease in the response rates in later years. In this respect, the relatively high response rate in year 4 is more atypical than the low response rate in year 3. This may be interpreted as being due to year 3 students focusing on orienting themselves within their new clinical attachments and not engaging with EEMeC to a great extent while year 4 students, already established in their attachments, were using EEMeC to research and submit coursework as well as to link back to their peers and tutors.

It is acknowledged that using EEMeC as the medium to evaluate itself could have introduced bias to the returns, in that only EEMeC users can have responded. However, the high response rates achieved, the near-mandatory requirement for students to access EEMeC regularly and the high commitment to evaluative activity across the course in general are considered to have ameliorated the effects of any such bias.

Analysis and interpretation

Triad analysis (effectiveness/utility/importance)

Questions 1–54 were delivered in triads framing the same question in terms of general effectiveness, personal utility and personal importance. A mean for each respondent for each of the three stem variants was taken. This was then analysed for internal reliability by calculating a Cronbach's alpha reliability coefficient, which showed a high degree of internal reliability ($\alpha=0.8652$). Non-parametric correlation analysis (Spearman's ρ) was then carried out between each of the three pairings (respondent means) with the following results:

- there was a significant positive correlation between general effectiveness and personal utility ($\rho=0.914$, $n=821$, $p<0.0005$). This indicates that there are fairly balanced feelings as regards EEMeC; respondents did not rate EEMeC as particularly good or bad in general relative to EEMeC's usefulness to them. Analysis of this correlation would be expected to indicate whether a VLE had a particular subjective reputation-bias relative to its objective evaluation. In the case of EEMeC effectiveness and utility were essentially equivalent in the respondents' minds.
- there was a significant positive correlation between general effectiveness and personal importance ($\rho=0.429$, $n=818$, $p<0.0005$) and a significant positive correlation between personal utility and personal importance ($\rho=0.448$, $n=818$, $p<0.0005$). Respondents considered EEMeC's importance to them was less than its perceived general effectiveness or personal utility. These pairings are taken to indicate the degree to which the VLE is the medium for course business. High positive correlations would indicate that the VLE was the principal medium for the course, no correlation that the VLE was no more or less important than other media for the course and high negative correlations that the VLE was of little or no relevance to the course. The data reflects the situation that EEMeC is a significant medium for course business but not the largest or most important one.

Table 4. Factor analysis and interpretations (extraction method: principal component analysis, rotation method: varimax with Kaiser normalization)

Factor	Rotation Sums of Squared Loadings			Questions with eigenvalues over 0.3	Interpreted factor description
	Total	% of Variance	Cumulative %		
1	9.296	15.756	15.756	3,6,9,12,15,18,21,24,27,30,33,36,39,42,45,48,51,55,56,57,58	Personal importance and relevance
2	3.875	6.567	22.323	1,2,7,8,25,26,46,47	General course participation
3	3.781	6.409	28.732	13,14,15,34,35,37,38,43,44	External connectivity
4	3.589	6.083	34.816	17,22,23,28,34,35,41	Support of activities
5	3.380	5.729	40.544	17,31,32,40,41,42,43,44,45,46	Educational support
6	3.109	5.270	45.814	26,28,29,31,32,59	Personal logistics
7	2.204	3.735	49.549	49,50,52,59	Feedback
8	2.187	3.707	53.256	19,20,25	Authority
9	1.978	3.352	56.608	45,52,53,54	Tracking and protection
10	1.906	3.230	59.838	4,5	Communication
11	1.887	3.199	63.037	16,17,52	Assessment
12	1.848	3.132	66.169	2,10,11	Provision of Information
13	1.720	2.915	69.084	25,46,47,59	General support

LAF validation

Since the questions had originally been generated from the LAF, it was important to verify that the predicted mapping between questions and the LAF was statistically valid. Factor analysis and inter-item reliability and correlation tests were performed.

The factor analysis identified fourteen underlying significant factors with eigenvalues higher than 1.0, the first 13 of which were interpretable with the dominant factor being that of 'personal importance' (these are shown in Table 4). Although the factor analysis was interesting there was no strong equivalence between it and

Table 5. Questionnaire-to-LAF map and inter-item reliability

Question triad	Mutuality	Competence	Continuity	Orientation	Reflection	Exploration	Convergence	Coordination	Jurisdiction
4, 5, 6	x	x	x					x	
7, 8, 9	x	x	x			x			
10, 11, 12		x	x	x			x	x	
13, 14, 15	x				x			x	
16, 17, 18		x		x				x	x
19,20, 21		x	x				x	x	x
22, 23, 24	x	x	x						
25, 26, 27			x	x			x		
28, 29, 30			x	x	x				
31, 32, 33			x	x	x				
34, 35, 36			x	x				x	
37, 38, 39		x		x		x			
40, 41, 42		x		x	x	x			
43, 44, 45			x	x			x	x	x
46, 47, 48	x	x					x	x	
49, 50, 51	x		x					x	
52, 53,54		x	x						x
Cronbach's inter-item alpha reliability coefficient	0.886	0.917	0.926	0.920	0.838	0.818	0.876	0.914	0.823

the LAF map although there was some congruence. However, as the responses were overall very positive and therefore heavily negatively skewed, there was a low level of variance and therefore a factor analysis would not be expected to be a particularly illuminative tool.

An inter-item test of reliability (Cronbach's alpha) was performed for each set of mapped responses to the LAF (as shown in Table 5). This was performed on the personal utility component of each triad (the previous section displayed that there was a very high correlation between effectiveness and utility and a reasonably high correlation between utility and importance). The results show a strong level of consistency across the question groups and therefore an acceptable level of reliability for the question-LAF map (none of the reliability coefficients were <0.8).

A non-parametric correlation analysis (Spearman's ρ) was also carried out for all item pairs. There was no significant difference between the mean correlation for the LAF mappings and the overall mean correlation. This indicates that, although a reasonable level of reliability has been established, the overall mapping is not very strong and further work needs to be done to refine this part of the instrument.

Overall LAF analysis

Having established the question-LAF map, each factor was analysed for each respondent group. The results of this are shown in Figure 1. From this a ranking of factors was plotted as shown in Figure 2.

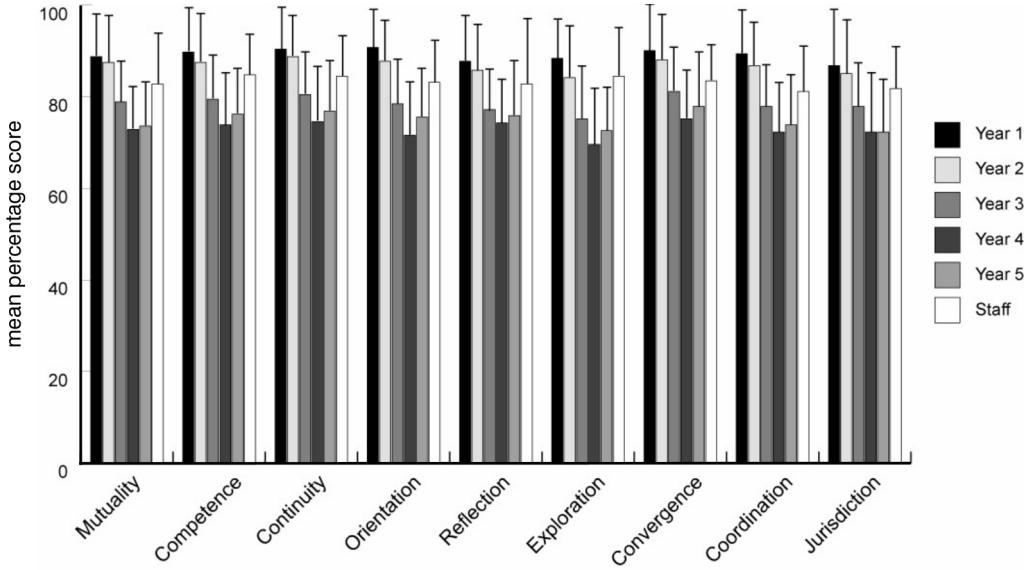


Figure 1. Learning architecture framework scores for EEMeC

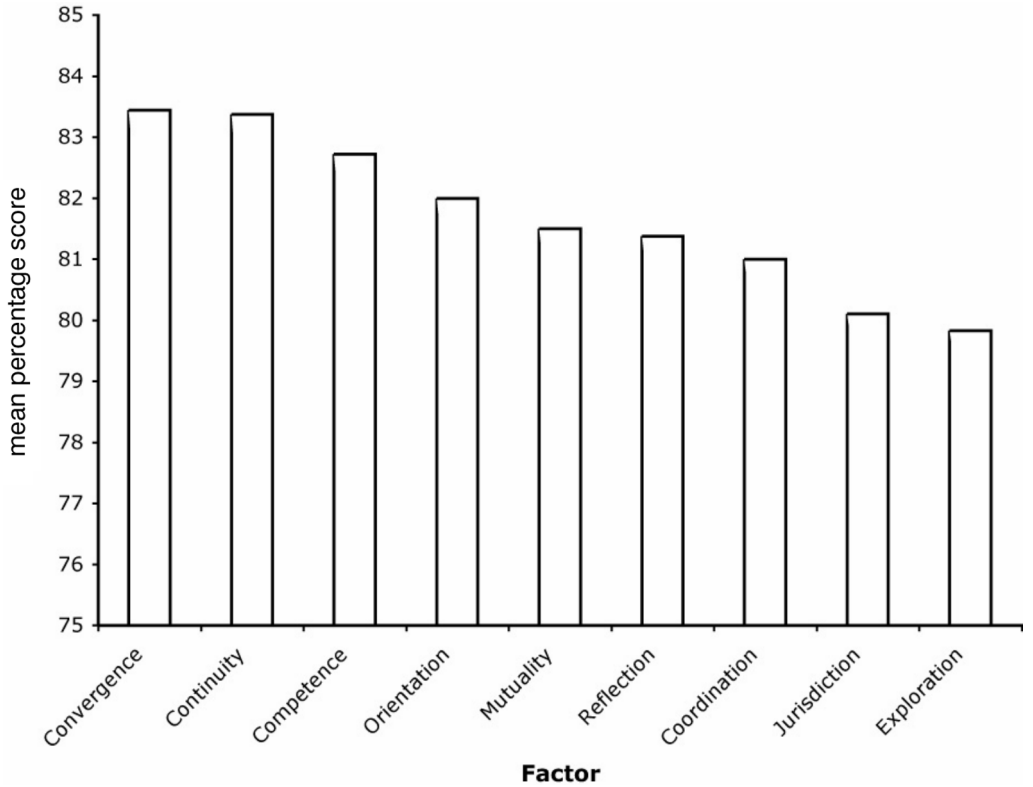


Figure 2. Ranked learning architecture framework factors for EEMeC

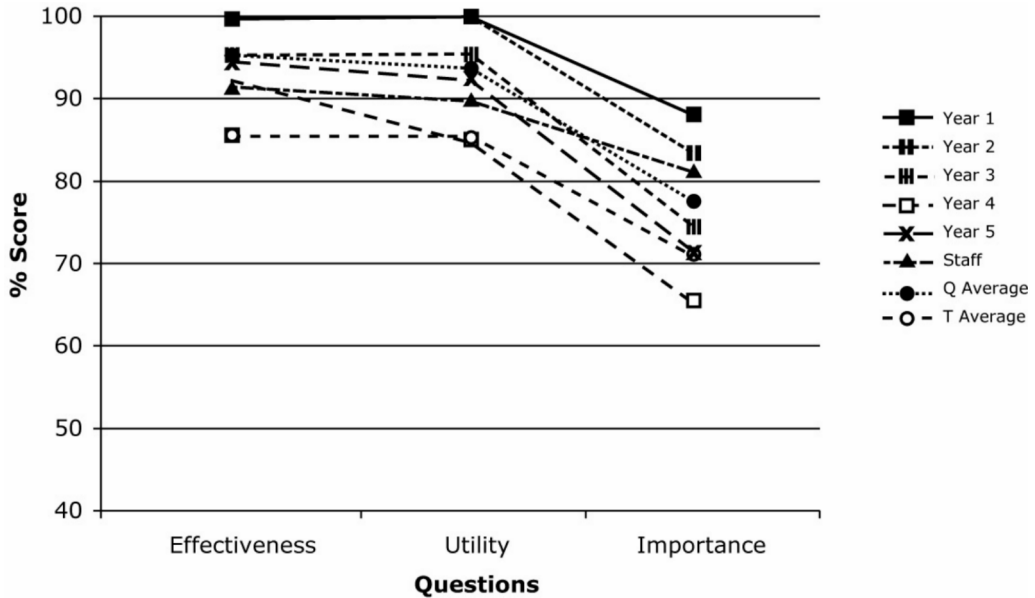


Figure 3. Involvement with the MBChB course: triad average SD = 0.17, overall average SD = 0.30). Interpretation: EEMeC is considered to provide a highly effective and useful service to the MBChB community of practice in engaging it with the course. However, because the course is not predominantly delivered or mediated online, EEMeC is considered to be less *important* by the community than it is *effective* or *useful*. Although there is deviation between cohorts, this is relatively small and indicates a relative consensus response. The staff cohort shows atypically higher importance rating while year 4 shows an atypically low importance rating. Conclusion: EEMeC is successfully mediating members of the MBChB community engagement with the course. There is a gradual decrease in score across the course with staff scoring somewhere in the middle.

Individual question triad analysis

Each of the question triads was then analysed across respondent groups and in comparison to the triad average and the overall average. A graph for each triad provides a useful illustration of the dynamics within the course community for each of the 18 issues addressed. A few examples are shown in Figures 3 and 4.

Free-text analysis

Question 60 was a free-text response allowing any free-text comments to be added. About 30% of respondents took up this option. Their comments largely echoed the findings from the rest of the questionnaire or made suggestions for specific changes or developments in the system. These were fed back to the course community as the basis for discussion within EEMeC’s steering and user groups. These comments will also contribute to the overall triangulated VLE evaluation but fall outside the immediate scope of this paper.

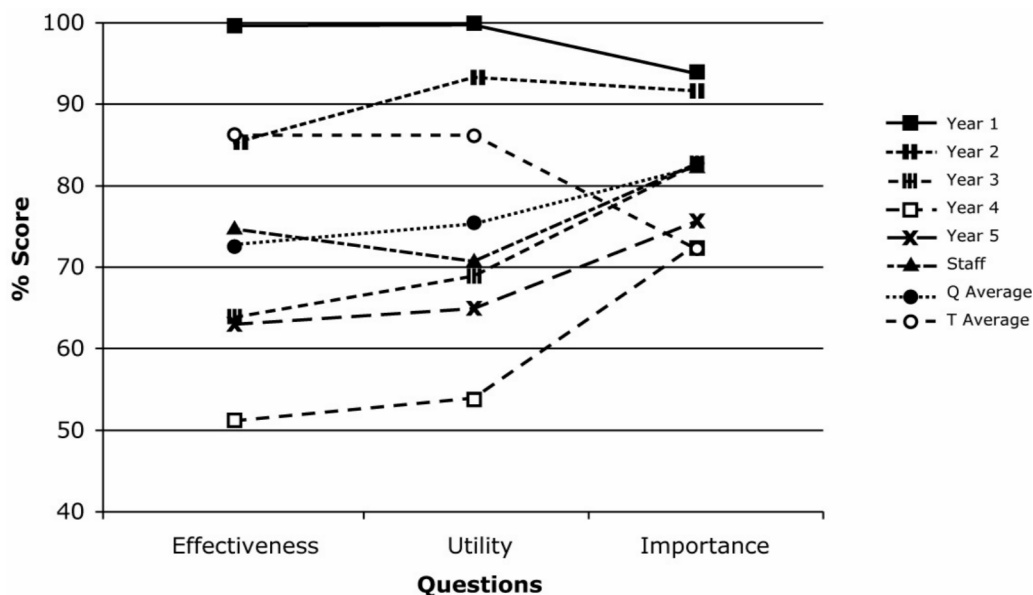


Figure 4. Providing timetabling and scheduling information: question average $SD=0.43$, overall average $SD=0.30$. Possible interpretation: EEMeC's provision of timetabling and scheduling information, although scoring reasonably well, shows little consensus between cohorts. Most rated the *importance* of this factor higher than EEMeC's *effectiveness* and *utility* in this area. *Personal utility* was generally rated higher than *general effectiveness*. Conclusion: although not entirely lacking, EEMeC's provision of timetabling and scheduling information to the MBChB community could be significantly better. This appears to be a priority for the community that currently considers itself to be insufficiently well supported.

Longitudinal and parallel measures

The Edinburgh MBChB is heavily evaluated (like most other medical courses). Evaluation fatigue is a major concern and, as the instrument is relatively large, a repeat survey will not take place until the autumn of 2004 (establishing an 18 month cycle).

This evaluation has recently been used on a VLE supporting undergraduate medical education in another university. A full data analysis is not yet available but provisional analysis indicates a similar ranking of LAF factors to those shown in Figure 2.

Respondent group bias

No accommodation has been made in the analysis for cohort bias as each academic year in the course has the approximately the same make up regarding gender, ethnicity etc and there is a significant mixing up between cohorts mid-course as about 40% of each year's students take a year out to do an intercalated honours course.

Delivery bias

As the VLE was the medium for delivery of the questionnaire, it is reasonable to anticipate a degree of bias from the fact that EEMeC users were the only ones who could have responded. The fact that use of EEMeC is mandatory for a number of course activities and the high student response rates indicate a fairly comprehensive coverage of the overall population however.

Discussion

The development and use of an evaluation instrument (based on Wenger's learning architecture framework) to investigate the utility of a VLE in the context of a specific course context has proved to be useful in capturing many of the dynamics of a VLE-in-use and thereby contributing to a broader holistic evaluation process. Validation of the instrument has however pointed out limitations in the mapping between the instrument and the LAF. Further work therefore needs to be done in refining this tool both in terms of the mapping between questionnaire items and the LAF and in carrying out longitudinal and parallel studies using this instrument.

The application of the evaluation instrument to an undergraduate medical VLE (EEMeC) indicates that there are aspects of the VLE that can be improved, particularly in the areas of course coordination (e.g. timetables), jurisdiction (e.g. rules and authority) and exploration (e.g. secondary learning materials) while other aspects are relatively strong. Analysis of its effectiveness, utility and importance show that the community using EEMeC has a reasonably realistic view of this VLE, and analysis of the individual questions provides feedback on specific aspects of how the system relates to its community of practice.

It is important to emphasise that this is a theory-based approach, which is predicated on a pre-existing course community of practice. In those situations where this is a valid assumption, for instance in subjects such as medicine, then it has immediate relevance and utility. For other situations, for instance in modular programmes of study, where communities of practice may not equate to a course (or even exist coherently at all) then there may be less relevance in such a study, although a module may in some cases retain a degree of internal coherence as a community of practice.

This approach to evaluating VLEs does not provide information on what the VLE can do, nor what its features are or even how it is used. These can be obtained from other sources and indeed these are usually already reasonably well known within the VLE's user community. What it does help to provide is a perspective of how successfully the VLE is serving the communities of practice involved with the course in question, and thereby is able to provide pointers to areas in which it could be improved to the benefit of that community.

It is important to restate that this approach is intended to contribute to a multidimensional 'triangulated' approach to VLE evaluation; other components may include log file analysis, use case expressions and analysis of impact although elaboration of these approaches are outwith the scope of this paper.

The proposal is that the LAF is not presented as a prescriptive framework around which a VLE should be designed and built; such prescriptive frameworks have been identified as inimical to professional practice (Lisewski & Joyce, 2003). It is descriptive rather than prescriptive; the themes used and the insights gained are recommended as tools to inform the reflective practices of those responsible for the design and delivery of VLE systems within coherent community of practice contexts. As Schwen and Hara point out:

a rich descriptive theory is not a warrant or recipe for the construction of certain phenomena, and a useful prescriptive theory may not provide a full understanding of the phenomena but rather a perspective on the conditions or circumstances of its applied use. (Schwen & Hara, 2003, pp. 261–262)

EEMeC has been developed organically over a number of years in response to the course's needs and wishes. This process was not directly based on the LAF although it was developed with the intent to support the whole 5-year programme. The LAF evaluation is therefore not intended to be a formal validation of a particular VLE design but rather as a contribution to the ongoing and iterative development of a VLE in support of its course and course community. It is also intended as a technique for practitioner-researchers who need immediate and reflexive insights into VLE design, deployment and support for specific educational contexts.

Conclusion

As Oliver (2000) observes: 'evaluation forms a unique meeting point between policy, theory and practice, and as a consequence, it seems unlikely that its practice will ever be uncontentious'. However, the approach presented here has been grounded in theory, is based around a holistic view of course-VLE instances and has provided significant utility to the authors in the evaluation of their own work. Oliver (2001) identifies a potential weakness in this kind of approach when others seek to use it—'the purpose of theory may not be fixed but may depend on the way users appropriate it'. Thus the level to which other users may find utility in this work may depend on the degree of agreement and alignment in approach and philosophy with that of the authors, and the contexts they are working in.

In situations where VLEs are used in programme settings where there is a strong and coherent community of practice then there is particular benefit in carrying out this kind of evaluation. There is often a major investment in, and dependence on, VLEs in these contexts (Cook, 2001), yet little in the way of appropriate evaluation instrumentation to supply those making this investment with feedback as to their efficacy and outcome.

It is hoped that the LAF framework can also provide a common language with which to compare different systems or ways of working with VLEs. Even if a VLE is only supposed to provide one or two aspects of support to a course it is still valid to apply this approach to investigate, for instance, the 'functionality creep' as the community using it tends to adopt originally unintended affordances from the system. All of these factors and issues will need further attention.

The development of approaches to evaluating VLEs as described in this paper can contribute insights on the multifactorial interactions between communities of practice and technology-mediated extensions to their learning environments and make a strong and valid contribution to broader holistic and triangulated evaluation programs.

Notes

1. <http://www.chest.ac.uk/datasets/vle/checklist.html>
2. <http://www.edutools.info/course/compare/all.jsp>
3. EEMeC-online at www.eemec.med.ed.ac.uk

References

- Alsop, G. & Tompsett, C. (2002) Grounded theory as an approach to studying students' use of learning management systems, *ALT-J*, 10(2), 63–76.
- Barajas, M. & Owen, M. (2000) Implementing virtual learning environments: looking for holistic approach, *Educational Technology and Society*, 3(3). Available online at: http://ifets.ieee.org/periodical/vol_3_2000/barajas.html/
- Breen, R., Jenkins, A., Lindsay, R., & Smith, P. (1998) Insights through triangulation: combining research methods to enhance the evaluation of IT based learning methods, in: M. Oliver (Ed.) *Innovation in the evaluation of learning technology* (London, University of North London).
- Britain, S. & Liber, O. (1999) *A framework for the pedagogical evaluation of virtual learning environments*. *JTAP Report 41*. Available online at: <http://www.jtap.ac.uk/reports/htm/jtap-041.html>
- Bruce, B. C., Peyton, J. K. & Batson, T. W. (1993) *Electronic quills: a situated evaluation of using computers for writing in classrooms*. Available online at: <http://alexia.lis.uiuc.edu/~chip/pubs/Equills/siteeval/index.shtml>
- Chalk, P. D. (2001) *Learning software engineering in a community of practice—a case study*. *CAL2001 Conference, Warwick*. Available online at: <http://www.ics.ltn.ac.uk/pub/conf2001/papers/Chalk.htm>
- Cook, J. (2001) *The role of virtual learning environments in UK medical education*. *JTAP Report 623*. Available online at: <http://www.ltss.bris.ac.uk/jules/jtap-623.pdf>
- Ellaway, R., Dewhurst, D. & Cumming, A. (2003) Managing and supporting medical education with a virtual learning environment—the Edinburgh electronic medical curriculum, *Medical Teacher*, 25(4), 372–380.
- Graham, G. (1999) *The Internet: a philosophical enquiry* (London, Routledge).
- Jenkins, M., Browne, T. & Armitage, S. (2001) *Management and implementation of virtual learning environments: a UCISA funded survey UK, UCISA*. Available online at: <http://www.ucisa.ac.uk/groups/tlig/vle/VLEsurvey.pdf>
- JISC and UCISA (2003) *Managed learning environment activity in further and higher education in the UK* (prepared by SIRU (University of Brighton), Education for Change Ltd & The Research Partnership).
- Koper, R. (2000) *From change to renewal: educational technology foundations of electronic learning environments* (Open University of the Netherlands). Available online at: <http://eml.ou.nl/introduction/docs/koper-inaugural-address.pdf>
- Lave, J. & Wenger E. (1991) *Situated learning: legitimate peripheral participation* (Cambridge, Cambridge University Press).
- Lee, M. & Thompson, R. (1999) *Teaching at a distance: building a virtual learning environment* (JTAP 033) (JTAP, UK). Available online at: http://www.jisc.ac.uk/uploaded_documents/jtap-033.doc

- Lisewski, B. & Joyce, P. (2003) Examining the five-stage e-moderating model: designed and emergent practice in the learning technology profession, *ALT-J*, 11(1), 55–66.
- Notess, M. & Plaskoff, J. (2004) Preliminary heuristics for the design and evaluation of online communities of practice systems, *eLearn Magazine*.
- Oliver, M. (Ed.) (1998) *Innovation in the evaluation of learning technology* (London, University of North London).
- Oliver, M. (2000) An introduction to the evaluation of learning technology, *Educational Technology and Society*, 3(4). Available online at: http://ifets.ieee.org/periodical/vol_4_2000/intro.html/
- Oliver, M. (2001) *What's the purpose of theory in learning technology?* (ALT Special Interest Group for Theory and Learning Technology Positional Paper). Available online at: http://homepages.unl.ac.uk/%7Ecookj/alt_lt/Oliver.htm
- Richardson, J. A. & Turner, A. (2000) A large-scale 'local' evaluation of students' learning experiences using virtual learning environments, *Educational Technology and Society*, 3, 4. Available online at: http://ifets.gmd.de/periodical/vol_4_2000/richardson.html
- Rogers, J. (2000) Communities of practice: a framework for fostering coherence in virtual learning communities, *Educational Technology and Society*, 3(3). Available online at: http://ifets.ieee.org/periodical/vol_3_2000/e01.html/
- Scarborough, H. & Corbett, J. M. (1992) *Technology and organization: power, meaning and design* (London, Routledge).
- Schwen, T. M. & Hara N. (2003) Community of practice: a metaphor for online design? *The Information Society*, 19, 257–270.
- Warren, P., Ellaway, R. & Evans, P. (2002) Meet George ... Using learning technology in a 'blended' approach to enhance the integration of knowledge and understanding across a 5-year medical course, *ASME 2002 Conference Proceedings*, Norwich, UK.
- Wenger, E. (1998) *Communities of practice: learning, meaning and identity* (Cambridge, Cambridge University Press).
- Wylde, K., Ellaway, R., Cumming, A. & Cameron, H. (2003) Electronic submission and delivery of student feedback, *AMEE 2003 Conference Proceedings*, Bern, Switzerland.