

Pragmatic meta-analytic studies: learning the lessons from naturalistic evaluations of multiple cases

Paul Lam*, Carmel McNaught and Kin-Fai Cheng

Centre for Learning Enhancement and Research (CLEAR), The Chinese University of Hong Kong, Hong Kong

This paper explores the concept of pragmatic meta-analytic studies in eLearning. Much educational technology literature focuses on developers and teachers describing and reflecting on their experiences. Few connections are made between these experiential ‘stories’. The data set is fragmented and offers few generalisable lessons. The field needs guidelines about what can be learnt from such single-case reports. The pragmatic meta-analytic studies described in this paper have two common aspects: (1) the cases are related in some way, and (2) the data are authentic, that is, the evaluations have followed a naturalistic approach. We suggest that examining a number of such cases is best done by a mixed-methods approach with an emphasis on qualitative strategies. In the paper, we overview 63 eLearning cases. Three main meta-analytic strategies were used: (1) meta-analysis of the perception of usefulness across all cases, (2) meta-analysis of recorded benefits and challenges across all cases, and (3) meta-analysis of smaller groups of cases where the learning design and/or use of technology are similar. This study indicated that in Hong Kong the basic and non-interactive eLearning strategies are often valued by students, while their perceptions of interactive strategies that are potentially more beneficial fluctuate. One possible explanation relates to the level of risk that teachers and students are willing to take in venturing into more innovative teaching and learning strategies.

Keywords: evaluation; meta-analysis; eLearning cases; risk-taking

Summarising lessons learnt in naturalistic studies

There are difficulties with the so-called ‘scientific research paradigm’ for education research. These include the complexity of interrelated factors in most educational settings; results from highly controlled scientific research designs not being replicable in real settings (Kember 2003); the difficulty of achieving matched groups in education; the ‘diffusion effect’ in controlled experimental designs in the same class or even institution whereby ‘control groups can be contaminated’ (Craven et al. 2001, 643) by the experience aspects of the treatment; and the well known ‘halo effect’ (Thorndike 1920) whereby students who know they are part of a treatment group perform better just because they think they are getting special treatment.

The naturalistic research paradigm thus seems better suited to educational research (Alexander and Hedberg 1994; Grasha 1990). Certainly, there is a large body of educational technology literature that focuses on developers and teachers describing and reflecting on their experiences with eLearning (Beetham 2005). Innovations are implemented in authentic educational settings and the collection of evaluation data is done in non-intrusive ways with

*Corresponding author. Email: paul.lam@cuhk.edu.hk

as little influence on the teaching and learning environment as possible. Naturalistic evaluation considers multiple perspectives and accepts that interpretation of interactions is not neutral and needs to be situated in a particular context of time and place (Froehlich 1994).

Naturalistic research studies typically involve a mixed-methods approach with an emphasis on qualitative strategies, including observation, interviews, focus groups, questionnaires, narratives, and video- and audio-tape recordings. Qualitative descriptions ‘provide a depth to student behavior that is largely absent from traditional measurement techniques’ (Grasha 1990, 108). In the same vein, Runciman (2002, 146) noted that ‘unconstrained by the need to reduce the data to a set of numbers, qualitative research may allow phenomena to be studied from more perspectives and in greater depth’. Hammersley (1997) noted that an overemphasis on quantitative methods is not justified in educational research. Similar sentiments are echoed by Oliver and Conole (2003) who advocated the use of multiple models in eLearning evaluation and research in order to avoid an overly narrow view on what constitutes evidence.

Secondly, multiple sources of data are often collected in a naturalistic study because obtaining accurate representations with qualitative data is often challenging. Qualitative data are often generated by small groups of students whose views may not represent the broader population (Hughes and DuMont 1993). Opinions collected and interpreted can be easily influenced by the researchers’ practice (Hague 1993). Triangulation of results through using multiple evaluation strategies is thus often used to build robust and more general interpretations of phenomena (Guba 1981).

Thirdly, naturalistic evaluation studies may differ considerably. Different contexts and objectives for using a particular innovation usually result in a different set of evaluation strategies and instruments. The use of multiple sources of qualitative data makes comparing such studies even more difficult.

Conceptual framework

Three terms are central to this paper – ‘evaluation’, ‘pragmatic’ and ‘meta-analysis’. In educational evaluation, we propose that the intention is to gather evidence about the effectiveness of some educational practice or invention. The criteria for ‘effectiveness’ can vary markedly – evidence of ‘an increase in convenience’ is likely to be different from evidence of ‘an enhancement of student learning’. This is somewhat different from theory-oriented research where the intention is to develop or confirm specific theoretical constructs or models. While there is a clear overlap between evaluation studies and more theory-oriented research, the potential variety in design of evaluation studies brings a ‘fuzziness’ to the field and results in the need for guidelines about what can be learnt from single-case reports and which led to use of the term ‘meta-analysis’. We sought analytic strategies that could allow us to look across the data from a number of existing studies to reveal new patterns.

In this study, our objective was to use the experience that already exists to find suggestions for practice that might be, to some extent, generalisable to other contexts, even though generalisability is not always the aim of evaluation and, indeed, is rarely so in single case studies.

The study was ‘pragmatic’ in the sense that we looked for flexible principles that could guide our interpretations of existing data sets.

The analysis focused on naturalistic studies because its aim is to inform practice. Qualitative data play an important role because of the rich information they contain. We propose that generalisability of findings across a number of individual studies is enhanced using the following logic:

- If a phenomenon is found in several studies, then the phenomenon is more likely to be a general one.
- If a simple general trend is not found, the meta-analysis across multiple cases may lead to the identification of some critical factors which can differentially affect success or failure in different contexts.

Methodology

Our meta-analytic study involved small-scale studies undertaken during the e3Learning (enrich, extend, evaluate learning; e3L) project, which assisted teachers in using web-assisted teaching and learning in Hong Kong universities. It offered a range of design, development and evaluation services, and conducted a wide range of single-case eLearning evaluations. For each subproject or case, a team was formed whereby teachers had access to educational and technical support. Full details of the design of this project are in James et al. (2003) and the project website (<http://e3learning.edc.polyu.edu.hk>). This paper develops the preliminary analysis in McNaught and Lam (2005a), involving the full e3L data set and introducing two additional meta-analytic methods to build our understanding of the issues operating to affect eLearning in Hong Kong universities.

A tailor-made evaluation design was adopted for each case in order to serve the best interests of the individual teachers and cases (Lam and McNaught 2004). The customised approach resulted in highly individualised evaluation plans, questions and instruments, contrasting sharply with generic evaluation methods where standardised evaluation instruments are used across a range of sites. The process of evaluation consists of several stages:

- working out the evaluation questions with the teacher;
- choosing appropriate strategies to use in answering these questions;
- developing an evaluation plan;
- collecting data using the strategies selected (likely strategies were surveys, focus-group meetings, expert review and usability sessions), analysing the data;
- reporting the evaluation results to teachers; and
- assisting them in making decisions about what to do next.

This mechanism provided useful feedback that informed not only the teachers involved in each subproject but also other teachers interested in using the strategies (Lam and McNaught 2008). Over a 30-month period, the e3L project supported the web development, implementation and evaluation of 70 subprojects designed to actively assist teaching and learning in courses. Sixty-three of these successfully reached the evaluation stage. The cases were very different in nature; Table 1 shows the range of web-assisted functions involved. Obviously, each of these cases involved usually more than one web-assisted function. We have described the functions in two ways.

The first is the nature of the **interactivity** involved. Broadly speaking, interactivity can be either between the learner and online materials and tasks, or between people using the communication functions. Column 1 shows the four classifications we used:

- low interactivity, basically accessing static information and information basic to the course (Low int.);
- students interacting with materials and tasks (Ss and ‘stuff’);
- interactions about specific information (Info int.); and

Table 1. The nature of the web-assisted functions used in e3L cases.

Interaction	Web-assisted function	Support, Content, Feedback, Discussion				No. of instances in all 70 cases	No. of instances in the 63 evaluated cases
		S	C	F	D		
Low interactivity (Low int.)	Course background	●	○	○	○	7	7
	Notes/ PowerPoints	○	●	○	○	10	10
	Required readings	○	●	○	○	3	3
	Past papers	○	●	○	○	1	1
	Project/assignment archives	○	●	○	○	1	1
	Learning tools such as glossaries	○	●	○	○	1	1
Students interacting with materials and tasks (Ss & 'stuff')	Learning resources/links	○	●	○	○	26	20
	Multimedia materials	○	●	○	○	20	15
	Self-assessment	○	○	●	○	9	8
	Grade-giving assessments	○	○	●	○	2	2
	Games	○	●	○	○	2	2
	Teacher-student (one-on-one) communication	○	○	○	●	7	7
Interactions about specific Information (Info int.)	Assignment submission	○	○	○	●	4	4
	Formal discussion forums	○	○	○	●	27	27
More public interaction space (Public)	Exhibition of assignments	○	●	○	○	3	3
	Creation and exhibition of multimedia projects	○	●	○	○	2	2
	Inter-cultural discussions (across countries)	○	○	○	●	2	2
	Total					127	115

- more public interaction space. In our e3L cases, the forums were used for: concept clarification, peer review, working on set cases, doing team assignments/projects, debates, socialising/building community, special topics/issues/news, and self-reflection (Public).

We also classified the **web-assisted functions** in terms of four broad areas of web functions:

- study management and skills support (Support);
- resources/content for learning (Content);
- assessment and feedback to students (Feedback); and
- communication between teacher and students (Discussion).

These classifications allowed the classification of eLearning strategies followed in the subprojects. Seventeen such strategies were identified, based on both the technology use and learning design. For example, ‘assignment submission’ and ‘exhibition of assignments’ were differentiated since the former are primarily assignment dropboxes while the latter relates to displaying students’ work, often in forums. It should be noted that this classification is pragmatic, since it is linked to web functionality rather than student activity. It is, therefore, different from student-focused classifications; for example Laurillard’s (2002) classifications of media types (attention, inquiry, discussion, practice and production) and Conole’s (2007) taxonomy of student activity (assimilative, information seeking and handling, modelling, discussion, experiential and productive).

Our match between these four web functions and the 17 web-assisted functions is shown in Table 1. Our 63 websites involved all the 17 web-assisted functions (Column 2), and we had 115 instances of these in action (Column 5).

Table 2 shows the diversity of the evaluation instruments used. These favoured self-reporting strategies such as surveys and focus-group interviews. As mentioned in McNaught and Lam (2005b), we had difficulty persuading teachers that data that directly address students’ performance are desirable. Hong Kong has a culture that equates evaluation to using student satisfaction questionnaires.

Three main strategies were used to compare and contrast the lessons learnt in our 63 cases:

- (1) meta-analysis of the perception of usefulness across all cases;
- (2) meta-analysis of recorded benefits and challenges across all cases; and
- (3) selecting cases that have employed similar strategies and doing a meta-analysis of smaller groups of cases.

Perception of usefulness meta-analysis

We used an iterative approach to rating each web-assisted function in the context of the course in which it was used. In general, the ratings were allocated to each function according to the following three criteria:

- (1) *Amount of use*: The amount of usage recorded concerning that particular function; for example, the number of postings in forums, the number of attempts of quizzes, or the frequency of access to a particular resource. This information usually came from site logs and student surveys.

Table 2. Diversity of the evaluation instruments used in the 63 cases.

	Surveys (students)									
	Course-start	Task-end	Mid-term	Course-end	Surveys (teachers)	Focus groups	Interviews	Web logs	Students' forum postings	Others*
Feb03–June03	0	6	1	6	6	2	0	2	2	0
Jul03–Aug03	0	1	0	3	4	2	1	2	1	0
Sept03–Jan04	1	1	2	15	22	4	3	11	6	3
Feb04–June04	1	0	0	14	14	10	0	15	3	1
Sept04–Jan05	7	0	2	21	17	8	1	17	1	5
Feb05–June05	0	0	0	2	0	0	0	0	0	1
Total	9	8	5	61	63	26	5	47	13	10

* 'Others' includes classroom observations, and analyses of students' examination and/or assignment performance.

- (2) *Quality*: The quality of materials; for example, the thread length and the quality of the ideas in the forum postings (see Lam and McNaught 2006b for examples); the coverage of quizzes, their cognitive demand and amount of feedback; the course coverage of the notes and learning resources; and the use of media in the learning resources. Judgements of quality involved two reviewers individually rating resources or analysing opinions using pre-agreed criteria and negotiating differences.
- (3) *Perception of users*: Opinions were collected from either teachers or students (informants) concerning the usefulness of the functions in supporting learning, either via a survey question or comments made in interviews. Sometimes, informants were asked to rank the functions on a site according to their usefulness or when they indicated that they wanted particular functions to stay when the course is run again.

After a significant amount of discussion, one score was assigned to each web-assisted function in each case. We used a 1–5 scale, with 1 being ‘not useful’ and 5 being ‘very useful’. The interaction between the e3L team was essential here as we were called on by our colleagues to justify our decisions for any given case. The judgement involved two reviewers reviewing the relevant evidence. There was also, of course, checking between the ratings given to a function across a variety of different course contexts. In this level of analysis, the perceived usefulness of all the 63 evaluated cases was compared and contrasted.

Meta-analysis of learning benefits and challenges

The second meta-analysis was carried out to identify the general strengths and weaknesses of the eLearning strategies, concentrating on the qualitative data. Comments concerning benefits and challenges were usually collected through open-ended questions in surveys, focus-group meetings or teachers’ reflections. In a few cases, students’ performance in assignments or examinations was also considered.

The meta-analysis method involved the research team going over the evaluation reports for each project and extracting the benefits and challenges mentioned in them.

Meta-analyses of cases with similar themes

The meta-analysis of selected cases permitted a closer look at the eLearning strategies and the factors essential to their success. Through concentrating on the individual eLearning strategies in detail, the analyses were able to provide insights and suggestions concerning their use and usefulness in more detail than the overall meta-analyses outlined above.

Findings

Perception of usefulness: basic content is favoured

Figure 1 shows a ranking of the measures of perceived usefulness for the 17 web-assisted functions. This is also shown in Table 3 to compare the ranking with our dual classification of the functions. The data indicate that the most appreciated features concentrate on the provision of content; even provision of simple, static content is more highly valued than interactive strategies (four out of the top 10 belong to ‘low int.’). This is in special contrast

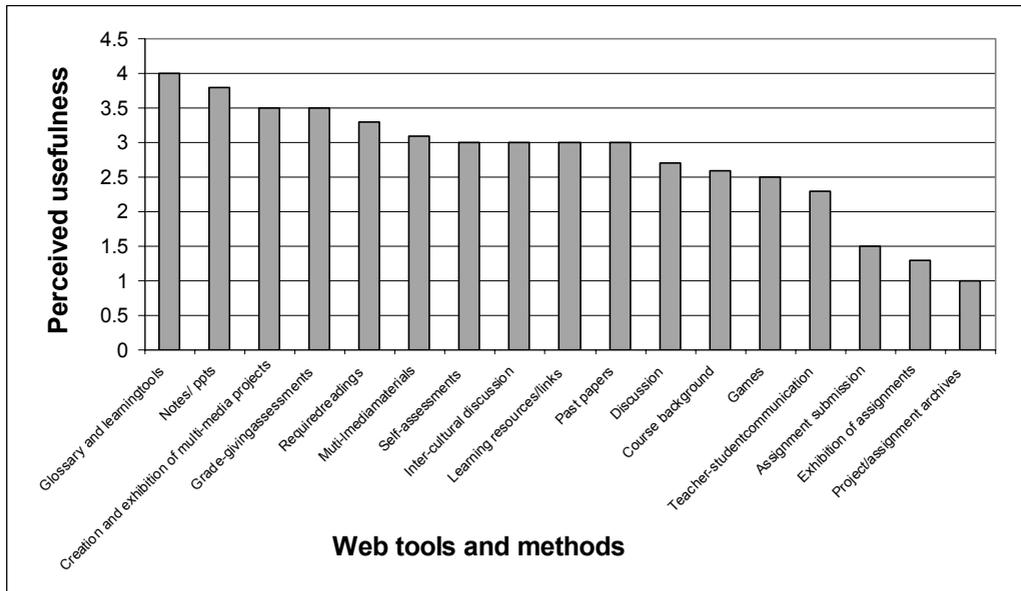


Figure 1. Ranking of perceived usefulness of the various web-assisted functions.

with those that require human to human communication ('info int.' and 'public'). The only exception may be the use of the 'creation and exhibition of multimedia projects' strategy, ranked third. This highly creative and student-centred activity made it very successful in both of the instances studied.

It is useful to look at the four 'top' ranked functions: learning tools such as glossaries, notes and PowerPoints, assessment tasks associated with grades, and creation and exhibition of multimedia projects. These all have rankings of 3.5 or higher on a five-point scale.

The highest ranked function is the provision of learning tools such as glossaries. One explanation is that, in Hong Kong, English is classified as a foreign language, and the standard of English is a matter of concern. It is not surprising that online dictionaries and glossaries of technical terms, especially with audio pronunciations or visual material, are highly valued. However, as universities in most western countries now cater for a wide international student mix, this finding may well be of broader interest.

The second ranking goes to notes and PowerPoint resources. Requiring low interactivity, and often highly didactic, these tools are, however, often treasured by teachers. It would seem that, although researchers advocate uses of the web that are interactive and seek to develop students' critical thinking, problem solving, information literacy, communication, etc., this will not be valued until the overall curriculum patterns in many universities change. Content-dominated curricula and a focus on end-of-course formal examinations with predictable patterns year to year are well supported by 'notes on the web'. Students, who understandably wish to succeed in the current system, wisely value these resources.

In the same vein, graded assessment tasks are valued. Without grades, assessment is less valued; for example, exhibition of assignments and project/assignment archives get the lowest rankings. However, positively, both instances in this group were good quiz exemplars, providing useful feedback to students. What is important to Hong Kong students is that they get marks for the work they do. The concept of working at quizzes in order to

Table 3. Ranking of perceived usefulness of the various web-assisted functions.

Rank	Interaction*	Web-assisted function	Support, Content, Feedback, Discussion				Average perceived usefulness (1-5 scale)
			S	C	F	D	
1	Low int.	Learning tools such as glossaries	○	●	○	○	4.0
2	Low int.	Notes/PowerPoints	○	●	○	○	3.8
3	Public	Creation and exhibition of multimedia projects	○	●	○	●	3.5
3	Ss & 'stuff'	Grade-giving assessments	○	○	●	○	3.5
5	Low int.	Required readings	○	●	○	○	3.3
6	Ss & 'stuff'	Multimedia materials	○	●	○	○	3.1
7	Ss & 'stuff'	Self-assessment	○	○	●	○	3.0
7	Low int.	Past papers	○	●	○	○	3.0
7	Ss & 'stuff'	Learning resources/links	○	●	○	○	3.0
7	Public	Inter-cultural discussions (across countries)	○	○	○	●	3.0
11	Public	Discussion	○	○	○	●	2.7
12	Low int.	Course background	●	○	○	○	2.6
13	Ss & 'stuff'	Games	○	●	●	○	2.5
14	Info int.	Teacher-student communication	○	○	○	●	2.3
15	Info int.	Assignment submission	○	○	○	●	1.5
16	Public	Exhibition of assignments	○	●	○	●	1.3
17	Low int.	Project/assignment archives	○	●	○	○	1.0

*The four types of interactivities:

Low int.: Low interactivity

Ss & 'stuff': Students interacting with materials and tasks

Info int.: Interactions about specific information

Public: More public interaction space

perform well in *other* assessments is not popular at all. It is a question of maximising the value from all time spent studying. Again, language may be a contributing factor: in many Hong Kong courses, the written materials are in English even when classes are in a hybrid English and Chinese mode. Students' reading speeds in a second language restrict the amount they wish to read, making them more strategic. Again, however, language challenges now exist in virtually all universities, and so this may have wider relevance.

The fourth highly ranked function is the creation and exhibition of multimedia projects. There were two instances of this in our case pool, one in public health and one in English. It is heartening that this function was seen to be effective. So, despite the comments made above about the overall didactic nature of higher education in Hong Kong, if course expectations are clear and oriented towards creative work, students will react accordingly.

Learning benefits: basic content is not favoured

Despite content-based eLearning strategies being very highly rated for perceived usefulness, the analysis on the level of reported benefits found that communication-rich strategies have great potential. Table 4 summarises the learning benefits mentioned in the evaluation

reports in the 63 cases. These remarks were further grouped under the four main types of eLearning strategies: low interactivity, students and ‘stuff’, information interactivity, and public (the first column of the table). The numbers in the cells indicate the number of times a benefit was recorded for a strategy in our cases.

There were many comments that the ‘low interactivity’ content was able to improve learning, ease downloading content, enhance problem-solving and analytical skills through the readings, and even change some of the students’ learning habits so that they are more ready to review and prepare for lessons. However, comments on these benefits were not as strongly or widely voiced as on more interactive types of content and communication strategies.

The interactive ‘Students and “stuff”’ type of strategies, for example, were considered to improve learning through revision and practice. With these strategies, the learning environment became more attractive and learning activities more varied. Multimedia materials in particular were considered to be excellent tools for students to view and review operational procedures. Self-directed learning was also enhanced; students remarked that they were motivated to learn by the new media.

The benefits of the teacher and student communication in the ‘information interactivity’ strategies were also impressive. Students mentioned better learning, increased interest in learning, and better communication with teachers. The benefits were not as strong, perhaps because there were fewer successfully implemented cases of this type in our study.

The benefits of the ‘public’ strategies were clear. Students learnt through seeing different people’s perspectives, and some through exchanging ideas with people in another culture. The learning environment was considered to be very engaging. Many types of learning skills were related to this by students: e.g. language skills (students had to express themselves clearly in communication), critical skills, argumentative skills and analytical skills. Lastly, attitudinal changes such as being more willing to cooperate were noted.

Challenges: basic content is safe

The analysis shows great differences among our eLearning strategy types. Table 5 illustrates the challenges mentioned in our cases, grouped under the four main eLearning strategies. The numbers in the cells are the frequencies of the challenges.

Interactive content and communication strategies were considered to have more challenges than the simple non-interactive strategies. While the ‘low interactivity’ strategies might fail technically (e.g. when downloading was problematic or slow) or in design (e.g. assigning too much reading for the students), in general the strategies were safe and successful.

The interactive ‘Students and “stuff”’ strategies, however, were demanding in all of the four areas identified. For example, in terms of perception challenges, students could treat the additional learning materials as unimportant and not pay attention to them. Technically speaking, more complicated materials led to problems with correct display on students’ computers, and there were database problems and data loss. The importance of a good interface for complicated tasks was also prominent. As for the learning content, more complaints were received about quality and quantity, since interactive materials are generally more difficult to create, resulting in fewer being developed and some not reaching high quality standards. Remember that seven (all ‘Students and “stuff”’ types) out of our original 70 cases failed to complete the evaluation phase. Lastly, there were concerns about out-of-class assessment; some students did not find online quizzes fair, particularly if they were used as part of the course grades.

Table 4. Benefits across the 63 cases.

	Low int. Total number of instances of this type: 23										Ss & 'stuff' Instances: 47				Info int. Instances: 11			Public Instances: 34	
	course background	notes/ ppts	required readings	past papers	project archive	learning tools	resources/ links	multimedia materials	self-assessment	grade-giving assessments	games	T-S communication	assignment submission	forums	exhibition of assignments	multimedia projects	inter-cultural discussions		
Knowledge																			
- Improved learning in general	2	7	1	1	1	1	18	8	9	3	2		8	4	1				
- Application of knowledge			1				1						1						
- Extended knowledge		1						3											
- Improved learning through communication											6								
- Increased revision and practice								4											
- Better memorisation								1	1										
- Better explanation through communication											2	1							
- Seeing different perspectives													4	2	1				
- Learning cultural difference																	2		
- Learning to quote external sources													1						
Environment																			
- Convenience		3																	
- Structured materials																			
- Visual attractiveness																			
- Flexible learning																			
- Variety of learning tasks																			
- More communication																			
- Engagement														4			2		

Table 5. (Continued).

	Low int. Instances: 23		Ss & 'stuff' Instances: 47				Info int. Instances: 11		Public Instances: 34					
	notes/ pts	required readings	project archives	resources/ links	multimedia materials	self- assessments	grade- giving assessments	games	T-S communication	assignment submission	forums	exhibition of assignments	multimedia projects	intercultural discussions
Content/materials														
- Insufficient materials				14	11	7	2							
- Insufficient quality				3	7	6								
- Insufficient instructions and guides				4	2	2								
- Inappropriate level of difficulty						1	1							
- Unhelpful discussion								2			1			
- Problem understanding peer work												3		
- Wrong concepts not clarified											2			
- Lacking instructions and guides												1		
Arrangement														
- No time to read		1												
- Poor layout design/organization				10	3									
- Infrequent update				4										
- Unfair assessment														
- Timing														
- Unclear arrangement														
- Poor arrangement of discussion														
- Heavy student workload												6		
											2	2		1

‘Information interactivity’ approaches could be difficult too. These were demanding in terms of teachers’ participation, responsiveness to questions, and the quality of feedback to students.

‘Public’ eLearning strategies were also difficult. Students lacked motivation to participate, resisted because they prefer face-to-face discussions, found some communication tools not to be user-friendly, and perceived a lack of teacher participation, especially when clarifications were needed. Tasks need clear and adequate guidelines to lead students through the activities. Even with a very good design, cooperative group work is usually time-consuming.

Thus ‘benefits and challenges’ analysis indicates that more interactive strategies are also considered more challenging to teachers and learners. However, they have the potential to deliver more learning benefits than the simpler, non-interactive strategies.

Lessons learnt from groups of similar cases

The third type of meta-analysis focuses on studies of cases using similar strategies. Four such analyses have been done, focusing on the challenging, interactive strategies. This allows the formulation of a list of the factors influencing the success of specific strategies across multiple contexts.

Lam and McNaught (2006a) examined three media-rich cases (a ‘Students and “stuff”’ strategy). The paper reports on the design and evaluation of the online courses and explores whether media-rich experiences can enrich learning at the full range of knowledge types and levels of cognitive reasoning outlined in the Revised Bloom’s taxonomy (Anderson and Krathwohl 2001). There was qualified support for benefits to student learning. The study also highlighted factors that influence success: attention to the quality and design of the media, considering student motivation, and focusing on feedback on learning during the course. Media and learning design thus are inextricably intertwined.

McNaught et al. (forthcoming) also studied more complex developments of the ‘Students and “stuff”’ type. It is rare to see failures in eLearning projects reported in the literature. The e3L cases that were not completed were examined to see why these were unsuccessful. The cases were: two sites with large numbers of high-quality video clips; two sites with flash animations that had to be scientifically very accurate; one detailed factory simulation; one role-play CD-Rom; and one portfolio system. Many challenges were uncovered. Each of the different stages (planning and development, implementation, dissemination, and evaluation) is critical and we analysed the challenges in each of these stages. Some challenges related to human factors – personal (e.g., workload) and interpersonal factors (mostly communication issues between teachers and development teams). These challenges can be compounded by external factors such as technological infrastructure, university policy and support, and the teaching and learning culture as a whole. One can view the development process as an intersection of a large number of factors that need to intersect in a mutual comfort zone. Sadly, this synergy does not always occur.

The third study addressed ‘public’ eLearning strategies, involving 13 reasonably active educational online forums (McNaught, Cheng, and Lam 2006). Data included student and teacher surveys, student focus groups, analysis of forum postings and examination of forum log data. Structured forums with a defined purpose generally had a higher quantity and quality of postings than free forums; student-centred fora (involving more student–student communication) were also of higher quality than discussions where communication was mostly from teacher to student. Further, analysing evaluation feedback from students and teachers identified three key factors that tend to affect forum success – ease of use, clear

facilitation and motivation to engage. The centrality of the role of the teacher was confirmed.

Lastly, Lam and McNaught (2006b) examined web-enabled communication in six cases involving peer- and group-assessment activities. These cases involved three web strategies – e-resources (using the web to deliver materials for discussion), e-display (displaying work for peers to critique) and e-communication (students discussing online). The data generally confirm that web-enabled peer- and group-assessment activities can produce positive results. However, these assessment activities need careful planning, as several factors affect outcomes. For example, the e-display function is not as popular as e-resources (in line with the first meta-analysis discussed above). Here, this was due to technical challenges: students will only upload work if it is easy to do so. Regarding e-communication, the importance of group dynamics and participation was highlighted. The skills needed to work effectively online are not the same as basic IT competency skills. In our cases both teachers and students were relative novices in online learning, since even students who have spent considerable time online may not be experienced in formal online learning situations.

Discussion

Table 6 summarises the main observations from the meta-analyses. Web functions that are basic and non-interactive, contrary to expectations, were perceived as more useful strategies than interactive ones. However, non-interactive strategies supported improvement in basic levels of knowledge, skills and attitudes, whereas other strategies were associated with learning at the higher levels. So, why would the pedagogically less beneficial strategies be more welcome?

High-risk, high-return

One explanation, highlighted in Table 6, concerns the high level of investment and risk teachers and students take in carrying out interactive strategies. As noted elsewhere, skills are needed in developing (Falconer 2007) and implementing online strategies (Moule 2007).

Table 6. Combined observations in the three meta-analysis methods.

	Perceived usefulness (meta-analysis 1)	Benefits (meta-analysis 2)	Challenges (meta-analysis 2)	Highlighted factors/issues (meta-analysis 3)
Low int.	Often high	Basic	Low	--
Ss & 'stuff'	Ranged between high and low	Can be very high	High	<ul style="list-style-type: none"> • Media quality • Learning design • Richness of feedback • Interaction with human, technology, environment and culture
Info int.	Ranged between high and low	Can be high	Medium	--
Public	Ranged between high and low	Can be very high	High	<ul style="list-style-type: none"> • Learning design • Group dynamics • Teacher and student participation • Support • Online learning skills

The fact that these strategies are challenging to develop and implement was confirmed in our third type of study, which examined more complex ‘Students and “stuff” and ‘Public’ strategies. Expected learning benefits were broadly confirmed in successful applications of these strategies, but various factors affect these outcomes such as the quality of the materials (particularly for multimedia), the design of activities, the readiness of teachers and students to participate and the support available.

Using a meta-analytic approach

The findings reveal that greater understanding can be achieved through comparing and contrasting naturalistic studies of eLearning use. Quantitative data may not always be directly comparable, as instruments and data vary. Qualitative data, on the other hand, yielded categories of description that could be analysed.

Pragmatic meta-analysis can also support generalisation across cases in terms of overall usage patterns, what works well and what does not (challenges), the strengths and weaknesses of a certain innovation and strategy, and factors influencing outcomes. This makes use of the many isolated and fragmented naturalistic studies that exist, and may help practitioners become better informed about the designs of innovations and eLearning strategies.

There are obvious limitations to this approach. As mentioned above, our naturalistic eLearning studies were heavily biased towards perception data rather than data on actual performance. Oliver and Conole (2003) called these “proxies for performance” and there is a tendency “to measure what is easily measured” (389). Generalisability of meta-analysis of this kind depends on both the quality and quantity of the individual studies available.

It is also interesting to note that pragmatic meta-analysis is not restricted to the study of completed or successful cases. As illustrated by our study of the seven incomplete cases, less satisfactory experiences can be very revealing.

Conclusion

We have suggested three pragmatic strategies for meta-analyses of naturalistic evaluations. Naturalistic eLearning studies vary in the evaluation techniques used, but the meta-analyses described in this paper can integrate findings and insights from studies with quite different designs.

This study indicates that basic and non-interactive eLearning strategies are often most welcome, while perceptions of more beneficial interactive strategies can fluctuate. This reflects that interactive strategies require support, and skills in development and implementation. It is a high-risk, high-return approach, and it seems that few teachers are currently willing to take this risk. However, several factors can interfere with these complex strategies. There are lessons for teachers embarking on an eLearning project, staff who support the use of technology in universities – both in terms of technical support and educational design – and those responsible for developing institutional policy about the effective use of technology in teaching and learning.

Acknowledgements

Funding from the University Grants Committee in Hong Kong is gratefully acknowledged, as is the collaborative support of colleagues in our three universities. All teachers generously agreed to their data being used in this paper.

References

- Alexander, S., and J. Hedberg. 1994. Evaluating technology-based learning: Which model? In *Multimedia in higher education: Designing for change in teaching and learning*, ed. K. Beattie, C. McNaught and S. Wills, 233–44. Amsterdam: Elsevier.
- Anderson, L.W., and D.R. Krathwohl. 2001. *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. Boston: Allyn & Bacon.
- Beetham, H. 2005. E-learning research: Emerging issues? *ALT-J* 13, no. 1: 81–9.
- Conole, G. 2007. Describing learning activities: Tools and resources to guide practice. In *Rethinking pedagogy for a digital age*, eds. H. Beetham and R. Sharpe, 81–91. New York: Routledge.
- Craven, R., H. Marsh, R. Debus, and U. Jayasinghe. 2001. Diffusion effects: Control group contamination threats to the validity of teacher-administered interventions. *Journal of Educational Psychology* 93, no. 3: 639–45.
- Falconer, I. 2007. Mediating between practitioner and developer communities: The Learning Activity Design in Education experience. *ALT-J* 15, no. 2: 155–70.
- Froehlich, T.J. 1994. Relevance reconsidered - towards an agenda for the 21st century: Introduction to special topic issue on relevance research. *Journal of the American Society for Information Science* 45, no. 3: 124–34.
- Grasha, T. 1990. The naturalistic approach to learning styles. *College Teaching* 38, no. 3: 106–13.
- Guba, E.G. 1981. Criteria for assessing the trustworthiness of naturalistic enquiries. *Educational Communication and Technology Journal* 29, no. 2: 75–91.
- Hague, P. 1993. *Interviewing*. London: Kogan Page.
- Hammersley, M. 1997. The relationship between qualitative and quantitative research: Paradigm loyalty versus methodological eclecticism. In *Handbook of qualitative research methods*, ed. J. Richardson, 159–74. Leicester: The British Psychological Society.
- Hughes, D., and K. DuMont. 1993. Using focus groups to facilitate culturally anchored research. *Culturally Anchored Methodology* 21, no. 6: 775–806.
- James, J., C. McNaught, J. Csete, P. Hodgson, and D. Vogel. 2003. From MegaWeb to e3Learning: A model of support for university academics to effectively use the Web for teaching and learning. In *ED-MEDIA 2003*, ed. D. Lassner and C. McNaught, 3303–10. Proceedings of the 15th annual World Conference on Educational Multimedia, Hypermedia & Telecommunications, June 23–28, in Honolulu, Hawaii, USA. Norfolk, VA: Association for the Advancement of Computers in Education.
- Kember, D. 2003. To control or not to control: The question of whether experimental designs are appropriate for evaluating teaching innovations in higher education. *Assessment and Evaluation in Higher Education* 28: 89–101.
- Lam, P., and C. McNaught. 2004. Evaluating educational websites: A system for multiple websites at multiple universities. In *ED-MEDIA 2004*, ed. L. Cantoni and C. McLoughlin, 1066–73. Proceedings of the 16th annual World Conference on Educational Multimedia, Hypermedia & Telecommunications, June 21–26, in Lugano, Switzerland. Norfolk, VA: Association for the Advancement of Computers in Education.
- . 2006a. Design and evaluation of online courses containing media-enhanced learning materials. *Educational Media International* 43, no. 3: 199–217.
- . 2006b. Evaluating designs for web assisted peer and group assessment. In *Self, peer, and group assessment in e-Learning*, ed. T. Roberts, 210–44. Hershey, PA: Information Science Publishing.
- . 2008. A three-layered cyclic model of eLearning development and evaluation. *Journal of Interactive Learning Research* 19, no. 3: 313–29.
- Laurillard, D. 2002. *Rethinking university teaching*. 2nd ed. London: Routledge.
- McNaught, C., and P. Lam. 2005a. Building an evaluation culture and evidence base for e-learning in three Hong Kong universities. *British Journal of Educational Technology* 36, no. 4: 599–614.
- . 2005b. What do teachers want to know about their student's eLearning? A study of 70 evaluation plans. In *Balance, fidelity, mobility. Maintaining the momentum?*, ed. H. Goss, 433–43. Proceedings of the 22nd annual Australian Society for Computers in Learning in Tertiary Education 2004 conference, December 4–7, at Queensland University of Technology, Brisbane. Brisbane, Queensland: Queensland University of Technology.
- McNaught, C., K.F. Cheng, and P. Lam. 2006. Developing evidence-based criteria for the design and use of online forums in higher education in Hong Kong. In *User-centered design of online*

- learning communities*, eds. N. Lambropoulos and P. Zaphiris, 161–84. Hershey, PA: Information Science Publishing.
- McNaught, C., P. Lam, K.F. Cheng, D.M. Kennedy, and J.B. Mohan. Forthcoming. Challenges in employing complex eLearning strategies in campus-based universities. *International Journal of Technology Enhanced Learning*.
- Moule, P. 2007. Challenging the five-stage model for e-learning: A new approach. *ALT-J* 15, no. 1: 37–50.
- Oliver, M., and Conole, G. 2003. Evidence-based practice and e-Learning in higher education: Can we and should we? *Research Papers in Education* 18, no. 4: 385–97.
- Runciman, W. 2002. Qualitative versus quantitative research - balancing cost, yield and feasibility. *Quality and Safety in Health Care* 11: 146–47. <http://qshc.bmj.com/cgi/reprint/11/2/146>.
- Thorndike, E.L. 1920. A constant error on psychological rating. *Journal of Applied Psychology* IV: 25–9.