



UCL E-LEARNING EVALUATION TOOLKIT

A framework for evaluating the pedagogical effectiveness of e-learning and technology-enhanced learning spaces

Vicki H.M. Dale, E-Learning Environments, UCL, July 2014

Contents

A framework for evaluating the pedagogical effectiveness of e-learning and technology-enhanced learning spaces.....	1
Introduction	3
Why evaluate?	3
Who should you consult?.....	5
What is it that you want to evaluate?	6
What type of evaluation do you want to conduct?	7
What questions should we be asking in an evaluation?.....	8
Technology-free educational designs	8
Technology-enhanced learning (TEL).....	9
Selecting an appropriate e-learning platform or tool.....	9
Evaluating the pedagogical benefits of an e-learning platform or tool	11
Technology-enhanced physical learning spaces	11
How do we conduct an evaluation?	12
The research process	12
Quantitative versus qualitative.....	12
Research questions and hypotheses.....	13
Research designs.....	14
A few words about epistemology and ontology.....	14
Post-intervention designs	14
Pre-post research designs.....	14
Test-control group research designs	14
Longitudinal research designs.....	15

Research methods	15
Questionnaires	16
Interviews.....	17
Focus groups and group interviews	17
Observation.....	18
System logs	18
Content/documentary analysis.....	18
Performance measures	19
Social network analysis	19
Sampling.....	19
Types of questions in surveys, interviews and focus groups.....	20
Closed questions	20
Open questions	22
Other interviewing questions and techniques.....	22
Quantitative data measurement	23
Nominal (categorical) data.....	23
Ordinal (categorical) data	23
Scale (continuous) data.....	23
Measures of central tendency and dispersion.....	24
Data analysis and reporting	24
Quantitative data analysis	24
Qualitative data analysis	25
Reporting the outcomes of qualitative analysis	26
Disseminating qualitative research findings	27
The ethics of evaluation research.....	28
Summary checklist	31
UCL e-learning case studies	33
ISD e-learning student survey.....	33
LectureTools pilot evaluation	33
@inDUCKtion evaluation	34
Blikbook evaluation.....	34
Making History group project (and MyPortfolio) evaluation	35
Glossary.....	36
References	38

Additional reading.....	39
Appendix 1: Example information sheet	40
Appendix 2: Example consent form	42
Creative commons licensing	43

Introduction

Evaluation is a critical part of educational development. Ideally, it should be considered at the outset of a project, so that strategies can be put in place to capture stakeholders’ expectations from the beginning, to be able to compare these with stakeholder experiences in practice. To be able to evaluate learning and teaching activities, whether face to face, blended or fully online, we need to consider the following:

- Why evaluate?
- Who should we consult?
- What is it that we want to evaluate?
- What type of evaluation do we want to carry out?
- What questions should we be asking?
- How do we conduct an evaluation?
 - The research process
 - Research designs
 - Research methods
 - Sampling
 - Quantitative data measurement
 - Types of interview questions
 - Data analysis and reporting
- The ethics of evaluation research

As well as considering these issues, the toolkit includes an appendix of case studies of e-learning evaluations carried out between 2013-14. These are brief overviews of the studies undertaken, the stakeholders involved, the research methods employed, and the general findings. A list of further reading material is also included.

Why evaluate?

In addition to justifying our learning and teaching strategies to senior stakeholders (in terms of efficiency and effectiveness), as educators or e-learning professionals, we have a duty of care to provide the best educational experience possible for our learners. Our undergraduate students in the UK now pay substantial annual fees, which has led to discussions about how we can put our ‘students at the heart of the system’ (Department of Business Innovation and Skills, 2011). Our students are likely to demand more enjoyable and pedagogically effective educational experiences that they consider to be meaningful and good value for money, as well as contributing to their employability. Postgraduate learners (often studying part time) and professionals paying for

continual professional development will have similar demands, though they are more likely to be concerned less with longer term employability and more with immediate application of knowledge in the workplace. As a knowledge society, we have a commitment to lifelong learning in general, which means evaluating the provision of education to the wider community. Finally, employers will need assurance that our graduates have received an education that is 'fit for purpose'.

CHECKLIST ✓

Why are you carrying out the evaluation?

To demonstrate stakeholder satisfaction (acceptability)	
To evidence a higher quality learning outcome	
To demonstrate that the teaching and learning activity/resource is 'fit for purpose' (Quality Assurance)	
To demonstrate value for money	
Other reason(s)	

Who should you consult?

To evaluate the quality of our educational provision, we may wish to survey or consult the following:

- **Learners** (undergraduate, postgraduate and lifelong) – our learners are the primary consumers of our teaching and learning efforts and are thus central to all teaching and learning evaluations.
- **Educators** (lecturers, professors, teaching fellows, teaching assistants) – you may wish to consult educators to ensure that the learning and teaching initiative is congruent with their educational aims, and to capture their experiences. Often the expectations of learners and educators will be very different, and ideally these should be aligned.
- **Teaching administrators** – TAs typically act as frontline support for students and educators, and act as a regular liaison between the two. Therefore, they provide a unique insight into the needs and experiences of both groups.
- **Learning technology/ e-learning professionals** – learning technologists will be able to advise on technology, learning design, usability, technical and functional considerations and will be able to reflect on the value of an e-learning resource in comparison with other e-resources.
- **Educationalists** – these experts in learning and teaching typically provide pedagogical support to educational initiatives, and will be able to comment on the likely educational value of proposed learning designs.
- **Management** (heads of e-learning, heads of department, heads of faculty, vice-provosts/principals and other senior management) – these stakeholders are likely to be most interested in whether educational interventions are fit for purpose and are cost-effective. They will make decisions with regards to investment of funds.
- **Employers** – these are the ultimate consumers of our 'products' in that they will be able to comment on the extent to which our graduates are fit for purpose.

You may be able to identify other stakeholders not mentioned here, such as professional bodies or societies. In addition, you may wish to consult other documentation, such as course handbooks. These will list the intended course or module outcomes, against which you can assess the quality of the educational intervention designed to support the attainment of these outcomes.

CHECKLIST ✓

Who/what will you be consulting for your evaluation?

Learners/course participants		Module/course leaders	
Educators (teaching fellows/ academics/ teaching assistants)		Teaching administrators	
Educationalists (learning and teaching advisors)		Managers	
Learning technology / e-learning professionals		Employers	
Other professionals e.g. library		Other	

What is it that you want to evaluate?

Evaluation strategies are typically used to help measure learner reactions; for example, do they like the resource? How confident do they feel as a result of using the resource? Evaluation may also help to measure whether learning has taken place; for example, are learners able to demonstrate that they have understood a particular concept, or that they know how to perform a particular skill? At a higher level, it may be possible to demonstrate whether learners have implemented what they have learned in practice, and at a higher level still, evaluation can help identify the result - whether learners’ new behaviours have had a demonstrable positive impact on their practice or employability. This is the hardest level to provide evidence for. These levels are presented in Kirkpatrick’s (1967) hierarchy of evaluation, reproduced below.

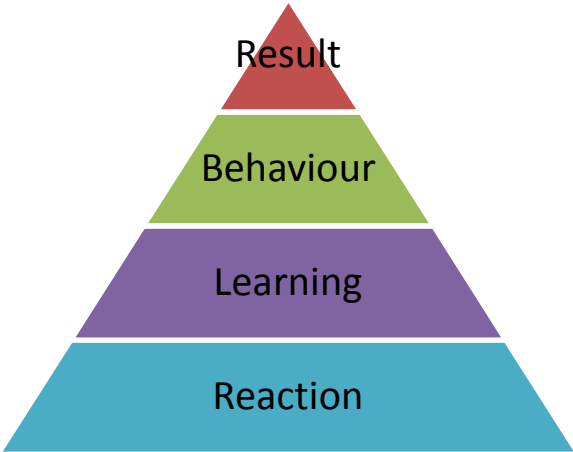


Figure 1: Levels of evaluation by Kirkpatrick (1967)

As an example, consider a teaching intervention which has been designed to encourage higher level thinking in undergraduates, such as a collaborative learning assignment designed to help them learn the material in a more effective way than traditional teaching, as well as encouraging the development of teamwork skills. At the first level, you might gauge learner opinion about the value of the exercise; do they like it; would they recommend it; do they perceive themselves to have

learned more about the topic? At the second level, you could measure learning through an assessment tool; how much do they know about their subject? Can they demonstrate evidence of teamwork? At the third level, behaviour, you would want to know if they are implementing what they have learned in practice; for example, are they working more collaboratively generally? The fourth level, result, might be measured by employer satisfaction.

Another example; suppose you want to evaluate the benefits of a new CPD course in restorative dentistry. At the first level, you would want to know if participants on the course liked the course, and if they found it useful, and you might even want to do a pre- and post- evaluation in relation to their confidence in particular techniques. At the second level, you might want to conduct pre- and post-tests to assess participants' knowledge gain or measure changes in their level of competence in demonstrating an individual clinical skill. At the third level, you would want to know if they had implemented any new techniques in practice; for example, is the dentist applying the new skill in practice with their patients? The top level, result, might be demonstrated by evidence that new treatments are more effective than traditional techniques in terms of being more comfortable for patients, or there being fewer post-operative complications.

Levels 3 and 4 typically require a longitudinal research design, to follow participants' behaviour changes over time as a result of the educational intervention. The problem with this approach is that it is virtually impossible to eliminate other confounding variables from the study, so most evaluation studies are based on level 1 and level 2 designs. What can be useful, particularly in the Continuing Professional Development (CPD) context, is to ask participants what they have learned, and how they have *applied* newly gained knowledge or skills in practice.

CHECKLIST ✓

Which of the following levels will you be seeking evidence for? (You can measure more than one.)

REACTION	Attitudes / perceived confidence	
LEARNING	Learning / assessed competence	
BEHAVIOUR	Behaviour change (skills in action)	
RESULT	Demonstrable impact	

What type of evaluation do you want to conduct?

There are three types (or levels) of evaluation that you may wish to conduct:

- Micro-level:** This would involve evaluating an innovation in the context of a small pilot study. In this case, we have typically not yet established the value of an intervention so there is an exploratory angle to this in terms of assessing the capabilities and limitations of the new resource and demonstrating proof of concept. An example might be trialling a single, new educational technology within an individual course. An example of this at UCL would be the [pilot evaluation study of LectureTools](#).

- **Meso-level:** This would involve evaluating the extension of a proven innovation on a wider scale after demonstrating a successful proof of concept. An example might be assessing the benefits and limitations observed when a proven technology is implemented on a wider scale; such as implementation of the new technology across a number of departments.
- **Macro level:** At this level, the intervention has been shown to be effective on a larger scale. The rationale for evaluating at this level would be to check that the intervention, now embedded as a core service, is fit for purpose. This might involve carrying out a quality assurance audit of existing educational provision at programme level, or even at an institutional level, as represented by the biennial UCL ISD student e-learning survey. It is also possible at this level that we are benchmarking our educational provision against specific pre-determined outcome measures, such as student satisfaction in the [National Student Survey](#) or the [Student Barometer](#). For a more comprehensive introduction to quality assurance and benchmarking of courses, see the QAA site (www.qaa.ac.uk).

CHECKLIST ✓

Which of the following level of evaluation will you be conducting?

Micro-level (proof of concept)	
Meso-level (wider implementation)	
Macro-level (quality assurance and enhancement)	

What questions should we be asking in an evaluation?

The sorts of questions we should be asking will depend on what type of educational intervention we are evaluating. Three types of pilot interventions are considered here:

- Technology-free educational designs
- Technology-enhanced learning (TEL)
- Technology-enhanced physical learning spaces (TEPLS)

In all of these situations, we can apply a framework based on the work of Ala-Mutka (2009), which seeks to identify: motivation, benefits, challenges, barriers and support, in conjunction with elements of Kirkpatrick’s hierarchy: reaction (level 1), learning (level 2), behaviour (level 3) and result (level 4). The questions listed below may be used as open questions in a survey or focus group, or they may be used to frame more closed, quantitative questions in a survey design.

Technology-free educational designs

In practice, it is unlikely that any educational designs will be entirely free from technology, as supporting materials for traditional teaching are likely to be made available on a virtual learning environment (VLE) or information about the classes may be disseminated via email, VLE or social media. This caveat notwithstanding, the sorts of questions you might want to ask of an innovative educational intervention would include:

- What is it that you want your learners to be able to achieve?
- What are the limitations of existing educational provision?
- Why are stakeholders motivated to pursue a new alternative?
- What perceived benefits does the new approach convey over existing approaches?
- What are the limitations or challenges associated with the new alternative?
- What barriers exist to the implementation or adoption of the new alternative?
- What additional support is required for the effective use of the new alternative?

This information might be self-reported through questionnaires or gathered using interviews or focus groups. It is also a good idea to physically observe students and educators in the face-to-face teaching situation.

Technology-enhanced learning (TEL)

Selecting an appropriate e-learning platform or tool

Although computer-aided learning has been available since the 1970s, the internet has opened up opportunities for increased technology-enhanced learning. In the first instance, a needs analysis should be conducted to determine the pedagogical requirements of a particular e-learning tool; what is it that the stakeholders are trying to achieve, educationally?

Following a consideration of pedagogical needs, the functionality of available technologies needs to be scrutinised. A feature comparison of available tools should be carried out, such as the one shown in Table 1 which shows a comparison of different online discussion tools. This includes questioning the extent to which the tool integrates with other e-learning systems and the availability of the tool on different devices and platforms, increasingly important as the Bring Your Own Device (BYOD) movement grows. The cost implications of different platforms should also be reviewed; for example, is there a one-off cost or an ongoing subscription? Many cloud-based tools have a freemium costing model – it is free for a limited number of users or functionality is limited in the free versus the premium version. Other considerations for choosing a learning technology tool or platform include the likely longevity of the tool (typically reflected in the size of the customer base or number of higher education institutions using the tool), the nature and amount of technical support offered, the user interface design and initial user experience (from aesthetic, functional and technical reliability perspectives). The technical stability of the platform is critical; this needs to be assured before additional feature comparisons and evaluations can be carried out.

	Blikbook	Knodium	Moodle	Piazza
Technical capabilities				
Anonymous postings	1	0	1	1
Ability to embed image in post	0	1	1	1
Mathematical equations	1	1	1	1
Chemical equations	0	1	0	0
Social sharing options	0	0	0	0
Customisable sharing options	0	1	1	1
Email notifications of posts/replies	1	1	1	1
Single sign-on	1	1	1	1
Grading of posts that integrate back to the Moodle Gradebook	0	0	1	0
Accessibility & Usability (for both app and website)				
Simple interface for navigating posts and posting new questions or replying to existing ones	1	1	1	1
Intuitive - can determine how to use it by clicking and reading on-screen prompts (without referring to help)	1	0	1	1
Contextualised help (if relevant)	0	1	1	1
Enable screen-reading software to navigate and read posts	?	?	?	?
All functionality available from a keyboard	?	?	?	?
Tab orders and jump-to menus to speed navigation using screen-readers	?	?	?	?
Accessible viewing modes, for example the ability to increase/decrease font size or change background and foreground colours or have a high contrast theme	0	0	0	0
Search within posts and subjects	1	0	0	1
Analytics features				
Enable staff to download analytics reports	0	0	1	1
Allow staff to easily query the log files / reports to see who has contributed	0	0	1	1
Cross-platform considerations				
App for Android	0	0	1	1
App for iOS	0	0	0	1
App for Windows mobile devices	0	0	0	0
App for Blackberry	0	0	0	0
Works on web browsers supported by UCL (I.e. Internet Explorer and Firefox) and for students without a smart phone	1	1	1	1
Reply by email				
Backups				
Automatically backs up posts and restores them in case of system failure	0	0	?	?
Integration				
Single sign on from Moodle Virtual Learning Environment (VLE)	1	0	1	0
LDAP or Shibboleth authentication otherwise (if not SSO above)	0	1	0	1
TOTALS	9	10	15	16

Table 1: Example of a feature comparison, in this case different online discussion/Q&A tools (credit: Matthew Lever)

Evaluating the pedagogical benefits of an e-learning platform or tool

Once an e-learning platform or tool has been adopted, evaluation questions you might want to ask include:

- How does the new technology functionally compare with alternative (traditional) tools?
- Why was the TEL innovation introduced? / Why are students and educators motivated to use it?
- Who is using the technology?
 - How often are people using the technology?
 - Do some learners use certain tools more than others?
 - Do people use the same tools in different ways?
- What particular features of the technology are they using?
- Why are they using it? What are the perceived benefits of the technology?
 - Pedagogically, technologically, organisationally?
 - Is the technology likely to have a demonstrable impact on performance at the individual, local or organisational level?
- What are the perceived challenges or limitations of the technology?
 - Pedagogically, technologically, organisationally?
 - Why are some users (students or educators) not engaging with the new technology?
 - Why are some particular features not being used?
- Are there any barriers to using the technology?
- Should the TEL innovation be continued? Should it be extended to other areas or expanded?
- What additional support do users require for effective TEL use?
- How might the learning design or technology be adapted or improved for future courses/cohorts?

This information might be self-reported through questionnaires or gathered using interviews or focus groups. It is also a good idea to physically observe use of TEL in a face-to-face teaching situation. Use of online learning can also be observed; for example, evidence of students' engagement with the TEL might be explicit (e.g. online discussion posts), or implicit (e.g. access logs available to course and system administrators).

Technology-enhanced physical learning spaces

When thinking about implementing a new technology-enhanced physical learning space (TEPLS), we would want to ask ourselves:

- What are the limitations of existing learning spaces?
- Why are stakeholders motivated to pursue a new alternative, either in terms of repurposing an existing space or designing a brand new learning space?
- What perceived benefits does the new space convey over existing spaces?
- What are the limitations or challenges associated with a new space?
- What barriers exist to the implementation or use of a new space?
- What additional support is required for the effective use of the new space?

Once a new learning space is in use, we would also want to ask:

- How are learners using the new space?
 - What do learners use the space for?
 - What do they physically do?
 - When do learners use the space? How often?
 - What is the perceived or actual impact on students' approaches to, and quality of, learning?
- How are educators using the space?
 - What do educators use the space for?
 - What do they physically do in the space?
 - When do educators use the space? How often?
 - What is the perceived or actual impact on the way that educators teach – logistically and pedagogically?

This information might be self-reported through questionnaires or gathered using interviews or focus groups. It is also a good idea to observe how the space is used over time in practice.

CHECKLIST ✓

Which of the following types of interventions will you be evaluating?

Technology-free educational designs	
Technology-enhanced learning (TEL)	
Technology-enhanced physical learning spaces (TEPLS)	

How do we conduct an evaluation?

The research process

Some familiarisation with the research process is required to be able to undertake an evaluation study. Any social sciences research methods textbook (e.g. Cohen, Manion et al. 2007, Robson 2011) will provide a comprehensive overview of the research process in education.

Quantitative versus qualitative

Very crudely put, there are two main research paradigms to be aware of; quantitative and qualitative. These have their roots in positivist (realist) and interpretivist schools of thought respectively. The positivist (quantitative) school of thought, which has its roots in the sciences, asserts that there is a single objective 'truth' in the world that can be observed and measured. The interpretivist (qualitative) school of thought, which owes its existence to the social sciences, is based around the concept that individuals have different, subjective perspectives and views of the world which represent equally valid 'truths'.

The distinction between quantitative and qualitative approaches in practice is not clear-cut; in designing quantitative studies, the researcher brings their own personal bias to the study in terms of

what is selected for study. In qualitative research, the accounts of individuals will be analysed, categorised and may even be summarised numerically (a quasi-statistical approach). Thus, there is an increasing move towards post-positivism which seeks to combine elements of both paradigms, and this is epitomised in the mixed-methods study. A mixed-methods approach might include a quantitative survey disseminated to a large study population which would seek to provide generalisable evidence, followed by interviews with a sample of participants, to try to gain insight into possible explanations for the survey findings.

CHECKLIST ✓

What approach will you take in your evaluation?

Quantitative (generalisable)	
Qualitative (insightful)	
Mixed methods (quantitative and qualitative)	

Research questions and hypotheses

Before conducting any research, including evaluation research, it is important to first develop a research question; this is what the research study seeks to find out. It might be to find out whether one educational approach is better than another in terms of student learning, or it might be purely exploratory in trying to find out student perceptions of a new TEL approach. Examples might be:

- What are students’ perceptions of a student-authored educational resource?
- What is the impact of a new classroom-based interactive technology on student performance?
- Does the use of regular, online assessments enhance students’ long-term retention of information?

The research question should be informed by prior reading and exploration around the topic under study. A literature review will provide information about what studies have been conducted previously, how they were conducted and what their findings were. This will help inform the new proposed study design, and ensure that the approach to the study is evidence-based, which will add credibility. Getting the research question right is important as this will influence the methodology and the specific methods used to capture the data.

Hypotheses are directional; they seek to prove or disprove whether a particular intervention has a certain effect; for example:

- Students with access to lecture recordings will perform better in examinations than students without access to lecture recordings.
- Students with access to a course-specific social study area will exhibit better teamworking skills than those without such a resource.

Hypotheses are not always necessary; they have their roots in the positivist (quantitative) school of thought; this is a ‘top-down’ approach where theory informs the data which will be collected. In a

very exploratory, qualitative study comprising observations of students, and interviews with individual students, the approach may be more of a 'bottom-up' approach where findings are said to 'emerge' from the data.

Research designs

A few words about epistemology and ontology

Philosophically, [ontology](#) represents the way we think about the world. [Epistemology](#) refers to thinking about knowledge – how is knowledge created and where does it come from? What are legitimate sources of knowledge? These beliefs and assumptions by the researcher shape their thinking about reality and how it is measured. Situated within a particular research paradigm, these philosophical assumptions will inform the research design, and methodologies and methods used. If an approach is epistemologically positioned then it means that the researcher has made particular assumptions about how knowledge is created. In *phenomenographic* studies; for example, the researcher assumes that knowledge is individually and personally created. There is no single 'truth' to be discovered and the researcher recognises that individual participants' experiences and beliefs are equally valid and seeks to report the variety of experiences. In *ethnographic* studies, the researcher is trying to understand a culture *in situ* as it operates as a whole. *Case study* research advocates that an individual or individual institution can be legitimately studied to provide insight into the workings of that individual or institution – no attempt is made to generalise the research findings to other people or institutions, but it can be assumed that the research can provide insight into other individuals in similar circumstances, or similar institutions. Typically, in evaluation research, you do not need to adopt a particular philosophical position, but it is important to be aware that how you approach your research will influence the way that you gather and analyse data.

Post-intervention designs

The most commonly used method for evaluating TEL is a post-test design, which simply involves gathering information about study participants – typically learners – *after* the educational intervention. How long after the intervention will depend on how important it is that the experience is 'fresh' in learners' minds, to avoid temporal bias, versus allowing sufficient time to allow the learners to apply what they have learned in practice. However, the danger of not capturing learner experiences immediately is that other extraneous factors may influence their perceptions and they may forget the specifics of the intervention that you are seeking to evaluate.

Pre-post research designs

Another frequently used research design is to measure stakeholder attitudes or performance immediately before and immediately after an educational intervention. As with other experimental designs, the longer the period of time between the pre- and post-assessments, the harder it is to attribute any changes to the specific intervention you are trying to evaluate. However, this can be a useful evaluation technique (e.g. see Dale et al. 2013 on [evaluation of a clinical logbook for promoting students' self-directed learning](#)). It is not only important to find out if a change has occurred; we need to understand *why* the change has occurred.

Test-control group research designs

Another commonly used method for evaluating learning technology, which owes its origins to the positivist paradigm, is that of a test group exposed to an intervention, which is compared with a control group which is not exposed to the intervention or which has some form of traditional

training instead. This is known as the randomised control trial (RCT). In such studies, subjects (participants) are usually given some sort of pre-test to assure the researcher that the participants in each group have the same level of pre-existing knowledge, or they are subsequently sorted into the groups to try to eliminate this bias. However, even controlling for such a factor, there are so many other confounding factors in educational research that it can be difficult to attribute outcomes to specific interventions. There is no guarantee that any increase in student satisfaction and/or learning in the test group can be attributed to the intervention. This type of research within the education community has attracted criticism for this reason (Morrison, 2001); however, it is still commonly used in science education research. Another difficulty with this research design is ethical; is it fair to give two groups of students an educationally different experience in the expectation that one group will outperform the other? For this reason many RCT experiments allow students to experience the alternative education provision, after their attitudes and performance have been measured for the purposes of an evaluation.

Longitudinal research designs

If you want to assess the long term impact of an educational intervention on learners' behaviour, a longitudinal study will be required. This might involve evaluating learners' experiences, knowledge or behaviour over a substantial period of time. For example, you might be interested to know if a student year group exposed to an intervention in their first year continue to behave or perform differently from the preceding cohort in subsequent years. As with the aforementioned study designs, it is important for the researcher to acknowledge any limitations in the study, such as being unable to isolate the research participants from being influenced by other unknown factors.

Research methods

There are a number of research methods that may be used to gather information for an evaluation study. One early guide, most of which is still very applicable today, is the [LTDI evaluation cookbook](#) (Harvey 1998). It briefly summarises the research process and provides an overview of commonly used research methods which may be used for evaluation purposes. These are summarised in Table 2. This list is not exhaustive; other methods such as the nominal group technique (see Varga-Atkins et al 2011) or the Delphi process may also be used in evaluation designs and are described briefly in the glossary. Readers are directed to the aforementioned research texts (Harvey 1998, Cohen, Manion et al. 2007, Robson 2011) for more information.

Method	Brief description	How administered
Questionnaire	Survey tool that may incorporate different types of questions (e.g. yes/no, multiple choice, Likert-type scale and open questions). Usually largely quantitative.	<ul style="list-style-type: none"> • Paper-based • Email-based • Internet-based
Interview	Structured, semi-structured or unstructured conversation with a single participant about their views and experiences.	<ul style="list-style-type: none"> • Face to face • Telephone • Skype
Focus group	Typically a semi-structured conversation with a small group of participants. Here, the participants are encouraged to reflect on each other's contributions.	<ul style="list-style-type: none"> • Face to face • Group Skype • Webinar • Virtual world

Group interview	Similar to a focus group, except that the interviewer is in sole control of the conversation; typically, each participant will be invited to provide a response to each question rather than engage in group discussion as they would in a focus group.	<ul style="list-style-type: none"> • As above
Observation	An ethnographic research method in which the researcher observes the participants. The researcher may be overt (visible to the group as an outsider) or covert (an insider within the group).	<ul style="list-style-type: none"> • Directly (in person) • Indirectly (via image, audio and/or video capture)
System logs	System logs (within Moodle, for example) will provide information about which students accessed which resources, when they accessed them and for how long. This method, when applied to large data sets forms the basis of learning analytics , where big data may be used to predict students' performance.	<ul style="list-style-type: none"> • Usually stored within the e-learning platform, or via website monitoring
Content/documentary analysis	Content analysis may be performed on a document or collection of documents produced by stakeholders. These may include personal memos or narratives, institutional or departmental strategies, course documentation, and may exist in a variety of formats (text, images, video).	<ul style="list-style-type: none"> • Content may be analysed quantitatively (e.g. word frequency) or qualitatively (to identify recurring categories and emerging themes)
Performance measures	Learners' performance data can be correlated with specific interventions to determine whether the intervention has had a significantly positive effect on performance.	<ul style="list-style-type: none"> • Learner grades may be stored in the VLE or within a separate system. Ethics permission is usually required to access individual assessment data.
Social network analysis	Analysis and visualisation of online network connections in social media.	<ul style="list-style-type: none"> • Tools such as Gephi or NodeXL

Table 2: Research methods that are typically used to generate or collect data in evaluation research

Questionnaires

Questionnaires are typically also referred to as surveys, although it is possible to 'survey' the views of participants using a range of methods. Questionnaires may be fully quantitative comprising closed questions or fully qualitative comprising open questions, but usually they are mostly quantitative with some open-ended questions. Questionnaires are used when a whole stakeholder group is to be surveyed, usually students but sometimes also educators and other stakeholders. Questionnaires may be performed pre- and post- an intervention, to determine whether exposure to the intervention has influenced participants' perceptions.

Questionnaires may be paper-based but are increasingly online, which can adversely affect response rates but which reduces the cost of consumables and manual data-entry. Again, the researcher needs to acknowledge any potential response biases in their report as a limitation of the study. A response rate of 30-40% is considered typical for online surveys. Nulty (2008) provides a useful overview of response rates in online versus paper-based surveys. Questionnaires may also be sent as an email, which is easy and ideal for very short questionnaires (up to five questions) because the intended participant can immediately reply.

Interviews

Interviews may be very structured – a standardised telephone interview comprising closed questions is one example of this. Conversely, they may be very unstructured ('conversational') and interviewee-led, with just a few prompts, where the researcher gains a richer insight into individual perspectives. These are typically carried out in phenomenographic/narrative studies, where participants provide insight into a topic of study by providing their own personal, subjective narratives. An example might be life history research in education. More typically in evaluation research, a semi-structured approach is used, where the interviewer has a series of limited standardised prompts but these can be followed up with other questions depending on the interviewee's responses. This allows the researcher to identify patterns across the participant responses more readily, while still allowing participants to contribute their ideas.

Conducting interviews face-to-face is best in terms of establishing a rapport with interviewees, particularly if the researcher has had little or limited previous contact with them, but this is not always possible. For this reason, some participants prefer a Skype interview where they can also see their interviewer on camera and vice versa, as opposed to a telephone interview. However, some interview participants may not feel uncomfortable 'on screen' and so may prefer an audio-only interview. Software such as Pamela (www.pamela.biz) can be installed to allow Skype interviews to be audio (and video) recorded.

Focus groups and group interviews

Advice varies about how many participants to include in a focus group discussion. There should be enough to enable participants to reflect on each other's contributions but not so many that the discussion is difficult to manage and direct. Advice in the literature varies as to the ideal number of participants; Kitzinger (1995) advocates 4-8 while Moore (2002) suggests 6-12; however, a group of 6-8 is probably optimal. As with interviews, focus groups will take up to 90 minutes of a participant's time, so it is usual to offer some sort of incentive or reward for participating, particularly for students; lunch or a book token is a suitable incentive/reward, although with the increasing demands on fee-paying students (and the danger of survey and focus group fatigue), it is becoming more difficult to attract voluntary participation with the offer of these incentives. As with interviews, the researcher should have several prompts or questions to open up the discussion, and be prepared to step in and redirect the discussion should it go off on a tangent. It is also important not to allow a focus group for evaluation research to descend into a 'moaning' session; there are other quality assurance forums and opportunities where students can vent their frustrations.

Another issue to consider is whether the teacher associated with the educational intervention sits in on a focus group discussion. Ideally, they should not be present, as students may be tempted to give

a more positive response to please their teacher than they would if speaking with an 'independent' researcher. This also removes participants' anonymity, which is an ethical issue.

Observation

Hearing from participants about their experiences of a new technology or technology-enhanced learning space is useful, but actually watching them use the technology or learning space in the course of their learning can provide a window on the learners and educators 'in situ'. The 'ethnographic approach' associated with observation has its roots in anthropology. The researcher may be an overt external observer; the advantage being that the observer can take an objective view. However, the presence of an outsider may influence the learners' behaviour (the [Hawthorne effect](#)). Alternatively, the researcher may be a covert insider; this will allow the observer to infiltrate the group under study, and to tap into group norms and behaviours, but this may make the researcher's observations more subjective. Another consideration when observing study participants is whether to observe them fully for a finite period, or whether to observe them periodically for shorter periods of time; this will largely depend on the logistics of the course, the researcher's availability and whether or not the researcher is interested in the experiences of the group over time.

One way of indirectly observing a learning space and gathering visual data about how the space is used over time would be to do a time-lapse video of the space. The Narrative Clip camera (<http://getnarrative.com/>) is an unobtrusive recording device that could be used to collect data about how students and educators use learning spaces; however, this presents a dilemma – ethically it is important to let individuals know that you will be photographing or recording them so that they can opt out if they do not want to be filmed. The difficulty this creates is that just with present observers, the study participants' behaviours may be skewed. For more guidance about evaluating technology enhanced learning spaces, see Roberts and Weaver (2006), Pearshouse, Bligh et al (2009) and Wilson and Randall (2012).

System logs

With online learning, system logs can provide information about what areas of content have been accessed, which activities have been engaged in, when and for how long users were logged in. Virtual learning environments such as Moodle can provide this information, although any exported data may need to be restructured to be analysed in a meaningful way. Such information is generally available to course and system administrators only. This ability to track users through online activities feeds into the concept of [learning analytics](#), where the educator can identify lack of engagement or poor performance in online tasks at an early stage, in order to intervene in a timely way and support students who are not engaging or performing well. A lack of engagement with online activities may not only reflect unmotivated or struggling students; this can also be a reflection of a problem with the technology or the underpinning learning design. However, it is important to also recognise that being logged onto a website and clicking links is not evidence of learning.

Content/documentary analysis

Any type of content or documents can be analysed; including:

- Personal narratives, such as written or video diaries, blogs, portfolio contents and their accompanying narratives
- Course documentation, such as course handbooks, timetables, syllabuses

- Strategic documentation, at departmental, institutional, national or even international levels
- Promotional material, such as prospectuses
- Social media content

Of course, the products of other research methods such as interview and focus group transcripts will also be subjected to a form of content analysis called ‘thematic analysis’ to identify categories of responses representing recurring responses and possibly emergent ‘themes’, depending on the extent to which the researcher is seeking to develop educational theory versus simply providing a descriptive summary of the perceived benefits and limitations of a particular intervention.

Performance measures

Performance measures allow the researcher to capture level 2 of Kirkpatrick’s hierarchy (learning); for example, has the intervention positively contributed to student learning? Performance measures may be obtained from online or traditional assessment results and are generally triangulated with use of, or perceptions about, a new TEL initiative. However, it is difficult to design an experimental study in education where research subjects and conditions are tightly controlled as they are in scientific research. There are many confounding factors which may influence student performance, and therefore any results should be interpreted with caution and the limitations of the study design made transparent. If an RCT research design is used, it is important to ensure that the two (or more) groups exhibit approximately equal performance in a pre-test.

Social network analysis

Social Network Analysis (SNA) can be used to get a sense of the dynamics within social media, the individual participants (‘nodes’ or ‘vertices’), and the interactions between them (the links known as ‘edges’). Free tools that may be used for this purpose include NodeXL (<http://nodexl.codeplex.com/>) and Gephi (<https://gephi.org/>). Data can be imported from social media (e.g. Twitter) directly or downloaded as an archive from the social platform and then later imported into SNA software. There are also third party tools such as Twitonomy (www.twitonomy.com) which can be used to summarise Twitter data (number of tweets, replies and mentions, as well as most active users, type of operating system, etc.). The @inDUCKtion project at UCL used Twitonomy to characterise the level of activity associated with this Twitter account.

Sampling

How many participants do you need to survey or interview? How representative do you need your sample to be? It depends whether you intend your findings to be generalizable to a wider population, or whether you only need to inquire insight into the subjective experiences of a particular group of individuals.

When planning a survey, intended to produce generalizable findings, it is worth doing a sample size calculation. This gives you an idea of how many responses you need from your survey population to generate representative results; that is, you can be 95 or 99% certain that the responses are representative. An example website that will allow you to do this is available online from [Creative Research Systems](#).

There are different sampling strategies to consider as well. For a full list of different sampling strategies see Cohen, Manion et al. (2007). Some of the most commonly used ones are listed in Table 3.

Sampling strategy	Description
Random	A random selection of participants from the whole population. This is likely to be more typical of quantitative study designs.
Stratified	Say for example, you needed to be sure of representation across different departments. You would need to consider the relative proportions of these departments in the population as a whole, and include the same proportion of participants from each department.
Snowball	Snowball sampling indicates that you start with a small number of participants, who in turn suggest other potential participants, and so on.
Convenience	A convenience sample is literally that; it is who you can get hold of at any one point in time. An example might be standing in a social area doing vox pops with as many students who are willing to speak with you.
Purposive	Here you are purposefully selecting particular individuals to represent particular groups or demographics; for example, a representative of each student year group.

Table 3: Commonly used sampling strategies in evaluation research

Types of questions in surveys, interviews and focus groups

Irrespective of the particular research methods employed, it is essential to have a knowledge of different question types, to facilitate an appropriate response from participants. The types of potential questions typically used in surveys are included in Table 4.

Type	Examples	Data
Dichotomous	Yes/No	Categorical / nominal
Multiple choice	Gender Faculty Which of the following (select all that apply)...?	Categorical / nominal
Likert scale	Strongly disagree - Strongly agree Not at all useful - Extremely useful Very poor - Very good Not at all confident - Extremely confident	Categorical / ordinal
Ranking	Rank the following (most to least important)	Categorical / ordinal
Numerical	Age	Continuous/ scale
Open-ended	What did you like best about ...?	Qualitative

Table 4: Types of questions

Closed questions

Categorical questions in surveys generally seek **closed** responses. In an interview, a Yes/No question is also an example of a closed question – there is usually only one of two responses possible; for example:

Have you previously used a VLE?

- Yes No

However, it is also a good idea to include a third 'Don't know' or 'Not sure' category for undecided respondents.

Multiple choice questions (MCQs) can invite a single response or multiple responses. If you are designing a paper-based questionnaire it is a good idea to use the conventions used in online surveys to distinguish between the two types of MCQ: use radio-buttons where respondents are being invited to select one option, and checkboxes where the respondents are being invited to select more than one option. With MCQs, it is always a good idea to include an 'Other' option, with a small text box below it to allow participants to enter their own data; this is particularly important for demographic questions since if the respondent does not see their particular category, they are likely to exit the survey, e.g.

Where do you live?

- England Scotland N. Ireland Wales
 Other (please specify)

Likert-type scale questions are very common to evaluation surveys, e.g.

How important is it to have training on digital professionalism?

- Very unimportant
 Unimportant
 Neither important nor unimportant
 Important
 Very important

There are a number of issues relating to Likert-type scales that are worth discussing:

- The length of the scale i.e. the number of options
- Whether to include a 'neutral' option or force respondents to make a choice
- The wording associated with the Likert-type scale

There are no definitive answers to any of these questions. A 1-to-5 scale is most commonly used, but the researcher may decide to select a broader scale, such as 1-to-7 or more, to allow for more variation in responses between different items. This is important if the researcher wants to be able to prioritise aspects of the learning design based on the evaluation outcomes, and/or conduct statistical comparisons between demographic groups. A 1-4 scale is sometimes used, particularly for course quality assurance forms; however, there is a tendency for respondents to ignore the extreme options, which essentially leaves the participant with a choice of two options! For this reason, if the researcher wishes the respondent to vote one way or the other, a 1-6 Likert scale is recommended, e.g.

How likely are you to recommend this course to other learners?

Very unlikely / Unlikely / Somewhat unlikely / Somewhat likely / Likely / Very likely

Many people prefer to include a 'neutral' option; however, this raises potential ambiguity – does the middle option reflect a 'middle' value, or is the respondent simply undecided? The wording of the items associated with the Likert scale can help to eliminate ambiguity. For example, if participants are being asked to rate their experience on a 'Very poor – Poor – Okay – Good – Very good' scale, then it is clear that the middle response is an average response. However, this is a very crude Likert-type scale. It is important to consider what construct you are actually seeking to measure, in relation to your original research question and the purpose of the survey. Vagias (2006) has produced a guide to potential Likert-type scale responses including measures of satisfaction, importance, agreement, priority, and frequency. However, while some of these can be considered roughly 'symmetrical' in the sense that there are the same number of negative and positive responses either side of the neutral option, some encourage a positive bias, for example:

Not at all familiar / Slightly familiar / Somewhat familiar / Moderately familiar / Extremely familiar

Another pitfall that the novice researcher needs to watch out for is combining constructs; for example 'Not very important – Not important – Useful – Very useful'. The construct useful also needs to be made explicit, for example 'How useful are the following for your learning'?

Open questions

With **open questions**, in a survey or interview, respondents can answer freely, they do not have to choose from a series of options. An open question would be:

Why would you recommend this course to a colleague?

However, the above question, on its own, is problematic. It is a **leading** question; it encourages the respondent to give a positive response. A better option might be to combine aspects of both questions:

Would you recommend this course to a colleague, and why or why not?

The danger with this approach is that you are asking **multiple** questions. Multiple questions should be avoided as respondents are likely to only answer one or the other. So, it might actually be better to have an initial closed question and invite respondents to explain their answer as follows:

Would you recommend this course to a colleague?

Why / Why not?

For more advice on question wording in surveys, see Cohen, Manion et al (2007) and Robson (2011).

Other interviewing questions and techniques

Other types of questions are useful when you are trying to encourage a response from interview or focus group participants. A **mirror** question indicates that you are listening and are seeking further information.

Interviewee: ... so overall I'd say it was a waste of time.

Interviewer: A waste of time?

A **probing** question can be used to elicit more detail from the interviewer:

Can you tell me a bit more about why it was a waste of time?

Alternatively, you can use a **prompt**:

Oh really?

In terms of interview techniques, it can be useful to consider models of clinical communication, such as the [Calgary-Cambridge model](#). This highlights the need to build a rapport with participants, the importance of body language as well as language, and the need to appropriately structure the consultation to gather the optimum amount of information. One particular technique to be aware of is the use of **silence** in interviews or focus groups. There is a temptation for novice interviewers to jump in too quickly with another question if participants do not immediately respond. However, these pauses are essential in giving your participants time to construct a response, and to allow them to speak freely without fear of interruption. Participants may take a few minutes to answer, but if you allow them to speak without interrupting, they will give you a considered and full response. They will typically look up and make eye contact with you again and/or nod when they are ready to answer another question. To ensure you get a really full response, you can use the 'Is there anything else...?' question to allow the participant to feel that they have answered the question fully.

Quantitative data measurement

In order to be able to analyse your data confidently, you need to know what type of data is being generated by your research.

Nominal (categorical) data

Closed survey questions, including dichotomous or MCQs, will result in categorical, **nominal** data.

Ordinal (categorical) data

The most commonly used ordinal data scale in surveys is a Likert-type scale response. Likert scale data results in categorical, **ordinal** data. Although the example given above has a qualifying descriptor for each option, some researchers omit interim descriptors e.g.

How likely are you to recommend this course to other learners?

Extremely unlikely ○○○○○○○○○○ Extremely likely

Note here that although the data collected is on a ten-point scale, the data is still ordinal (and categorical as opposed to continuous), because there is no indication of the actual numerical distance between the options. Summaries and statistical analyses of categorical data require non-parametric measurements and tests. However, it is generally accepted that if the scale is symmetrical that the mean may also be used as a measure of central tendency.

Scale (continuous) data

Scale or continuous data is relatively rare in surveys. Examples include:

How much money are you allocated for CPD, in Great British Pounds £.....

How many years have you been teaching undergraduate students? years

With scale (continuous) data, the units of measurement are proportional (£1000 is twice the amount of £500). This has implications for data analysis. Continuous data can be subject to parametric summaries and statistics. However, if one of the data variables is categorical (e.g. workplace grades) then non-parametric tests still have to be used.

Measures of central tendency and dispersion

Measures of central tendency and dispersion are useful to summarise numerical survey data. There are three measures of central tendency:

Measure	How calculated	Type of data
Mean	The average; the total score divided by the number of responses.	Continuous/parametric
Median	The middle value, when the responses are ranked from low to high.	Categorical/non-parametric
Mode	The most frequent response in the data set.	Categorical/non-parametric

Table 3: Measures of central tendency

Similarly, according to the type of data collected, there are different measures of dispersion (or spread) in the range of responses:

Measure	How calculated	Type of data
Standard deviation	The amount of variation in the responses in relation to the average value.	Continuous/parametric
Range	The maximum minus the minimum values.	Continuous/parametric or categorical/non-parametric
Inter-quartile range	The 75 th percentile or 3 rd quartile minus the 25 th percentile or 1 st quartile.	Categorical/non-parametric

Table 4: Measures of dispersion

Summarising numerical data as an average and a distribution is useful in summarising a single data set, or comparing the responses of different groups, or comparing ratings of different items.

Data analysis and reporting

Quantitative data analysis

Data should be summarised and visualised before statistical comparisons are made. Many people use Microsoft Excel to summarise and visualise their data, but this can also be done in SPSS, although before you can enter data in SPSS, you have to define your variables. Pallant's (2010) *SPSS Survival Guide* is an excellent introductory guide to using SPSS which explains the rationale behind different statistical tests and how to conduct them in SPSS and interpret the output. Another comprehensive text is Field's (2009 [ENREF 2](#)) *Discovering Statistics using SPSS*.

The type of data analyses you will conduct will depend on the type of data gathered. Usually, in evaluation research, you are dealing with quantitative data that is nominal (e.g. dichotomous variables such as yes/no or categorical data) or ordinal (e.g. Likert scale ratings) rather than continuous ('scale') data which is typically analysed using parametric statistics. Thus, you will use non-parametric tests to analyse your data. The commonly used tests are shown in Table 7.

Commonly used non-parametric tests	Equivalent parametric tests
Chi-square (2 x 2) or cross-tabs (2 x k) test – for between-group comparisons between two groups where the data is nominal (categorical)	Paired t-test
Mann-Whitney U test – for between-group comparisons between two independent groups where the data is ordinal (e.g. Likert scale data)	T-test for independent samples
Wilcoxon signed-ranks test – for between-group comparisons between two related groups (e.g. pre and post) where the data is ordinal (e.g. Likert scale data)	Paired t-test
Kruskal-Wallis Analysis of Variance – for between-group comparisons between more than two groups, followed by individual pairwise comparisons	Analysis of Variance with post-hoc tests
Spearman’s Rho – for correlation between two nominal or ordinal variables	Pearson’s correlation

Table 5: Commonly-used non-parametric tests

These tests will allow you to determine whether there is a statistical difference between groups or a correlation between two variables. The level of significance is generally indicated by $p < 0.05$. Researchers should be alert to the possibility of Type 1 or Type 2 errors; that is, rejecting the hypothesis when it is in fact true, or accepting it when it is false. To avoid the possibility of a Type 2 error, individual pairwise comparisons after a Kruskal-Wallis ANOVA should be subject to the Bonferroni correction; this means dividing the usual minimum p value of 0.05 by the number of pairwise tests being conducted, and accepting only p values less than this as significant. For example, if you are comparing the ratings provided by four different year groups (1v2, 1v3, 1v4, 2v3, 2v4, 3v4) then the minimum p value under which you can consider results significant is $0.05/6$, i.e. $p < 0.008$.

Qualitative data analysis

Qualitative data analysis can range from categorising short survey responses to making sense of extensive transcripts from interviews and focus groups. A commonly used method is thematic analysis, though as Braun and Clarke (2006) note, this term is used loosely in educational research. In practice, sections of text are ‘coded’ under different codes or labels. Usually these codes are then organised under categories. Typically, in evaluation research, you can then use that coding framework to provide a rich qualitative description of your findings.

It is possible to hand-code qualitative data by annotating print-outs of interview transcripts. If you only have a small data set such as four or five interview transcripts, or a couple of focus groups, it may not be worth investing the time in learning a program such as NVivo unless you are already familiar with it. However, with more complex or large data sets, including survey data, it is best to use NVivo or an alternative program as a way to code responses. This also allows you to easily retrieve the coded data as and when you need it for the purpose of reporting. The problem with immediately coding your data in NVivo, rather than reading through the transcripts or survey responses first, is the tendency to get bogged down in the details rather than viewing the data holistically. Thus it is always recommended that you take some time to read through the data to get a sense of familiarity with recurring themes. Some evaluation studies go one step further and look at the relationships between the categories as a basis for theory building. A useful introduction to more complex thematic analysis of interview transcripts is presented in a [YouTube video by Kent Löfgren](#).

Other qualitative coding approaches may be used in educational research, including [grounded theory](#), [interpretative phenomenological analysis](#), [template analysis](#), [framework analysis](#), [analytic induction](#), and others besides, depending on the researcher’s ontological and epistemological positions. However, it is unlikely that you will need to adopt any of these techniques in evaluation research.

Reporting the outcomes of qualitative analysis

Particularly with survey data, it is common practice to present the number of responses under each code. An example from the 2014 UCL ISD e-learning student survey is as follows:

Respondents were asked to state the three things that they like best about UCL’s IT provision. These are listed in the table below. Some of the individual text entries encompassed two or more aspects. Wifi on campus, support and Moodle were the most popular.

Service	n=	Service	n=
Wireless internet on campus	254	Library services including e-journals	74
Support (online, helpdesk, training)	228	Lecturecast	55
Moodle and online learning resources	208	Software including anti-virus	54
Comprehensive, easy, efficient, fast, free, reliable, up to date services	166	Accessibility and availability (of hardware, information and support)	52
Computers and cluster rooms	104	Remote access (WTS)	38
Email (Live@UCL)	97	UCL Go! app	32
Printing	80		

Alternatively, the results can be presented as follows, with illustrative quotes:

In the free text comments, students’ reasons cited for not bringing their tablet or e-reader onto campus included:

- Using e-readers for personal reading only (n=9)

“I bring my Kindle every day but not to support my studies, I read novels on it!”

- Not being able to view PDFs or textbooks properly (n=4)

“I don't use my e-reader because academic books and text books are usually not well formatted in Kindle format, so I prefer to just get the original books out the library, than buy a badly formatted book”

- Connectivity problems (n=4)

“My tablet does not allow me to connect to local printers and SharePoint, and therefore I have to carry both my laptop and tablet.”

Note that the above approach prioritises the findings based on the frequency of each code or category’s occurrence. It is not always necessary to quantify qualitative data in this way, particularly when working with smaller data sets. An alternative strategy is simply to provide a rich qualitative description of the findings. An example from an evaluation of Blikbook and comparable discussion tools is shown below:

Staff were motivated to use Blikbook or a comparable external tool (Knodium or Piazza) for a range of reasons. Moodle was perceived to be inferior to these tools in relation to its interface design (typically referred to as 'clunky'):

"I was motivated to give it a go in seeing whether – it's kind of a slightly more user friendly or modern looking interface than some of the Moodle forums – if it would attract more student use." (Blikbook user)

Moodle also did not have the ability to post questions anonymously (at the time that the external tools were adopted by educators) and it did not have a useful search function. Students did not appear to be engaging to any significant extent in Moodle discussions and the external tools were seen as another opportunity to engage them in discussion. Other features that staff were motivated to use included Blikbook's tie-in deal with publishers to distribute e-books (which has not yet materialised), and Knodium's support for ChemDoodle (a widely used platform for generating chemical formulae and graphics electronically) which makes it suitable for engineering students. Another perceived advantage of Blikbook was it that is an 'independent' system which – although it integrated well with Moodle – was sufficiently separate so that students were less likely to perceive that their participation in discussions was being monitored and contributing (unofficially) to their assessment results:

"I don't want any student ever, to feel – because of an adverse comment or positive comment they made – that they got disadvantaged or advantaged in some way or another ... I don't think it is the case in Moodle but I can quite see that when gossip starts getting around: 'Oh I raised this question and I only got 52 as well' ... students will make those connections." (Blikbook user)

Disseminating qualitative research findings

The results of evaluation studies may be reported in the following ways:

- **Twitter (microblog) or social bookmark** – for maximum exposure, include a link to publicly available research findings.
- **Blog post** – This is an informal way to promote good practice and showcase innovation in a public way. Blogs tend to be conversational in style and, in allowing readers to comment, can be used to stimulate online debate about the merits of specific educational approaches.
- **UCL Teaching and Learning Portal** – Case studies on the teaching and learning portal are a succinct way of promoting good practice publically. These tend to be journalistic in style, with links to useful resources. As with the blog posts, these can be used to encourage readers to engage in similar activities and help build a community of practice of educators and/or learners in relation to specific tools or educational approaches.
- **Conference presentation/poster** – Showcasing interim or summative evaluation findings through conferences and events can be a useful way to get direct feedback from colleagues. Writing a conference abstract and creating a narrative in the form of a presentation is a useful preliminary step to writing a paper, as it can be useful in crystallising thoughts about the study and presenting it in a clear, logical way.
- **Working paper or report** – A working paper or report may be a precursor to a submission to a peer-reviewed journal article or may be the end-point of the dissemination process. Reasons for only publishing your findings as a working paper include having a case study or results that are likely to be influential on other educators' teaching practice, but which might

be subject to methodological limitations such as a small sample size that might attract criticism from a reviewer and prevent the study being accepted for publication in a peer-reviewed journal.

- **Peer-reviewed journal article** – This is considered to be the most scholarly output of a research study.

UCL Discovery (<http://discovery.ucl.ac.uk/>) allows UCL staff to ‘publish’ conference posters and presentations, reports, working papers and the original text of peer-reviewed journal articles. There is a tension between the accessibility of research findings to the wider community of practice, and the perceived scholarlyness of a piece of work – journals have an ‘impact factor’ which will be of interest to academic colleagues who wish their work to be recognised by the Research Excellence Framework (REF). However, unless you publish in an open access journal, access to your work will be restricted to institutions with a library subscription for that journal. It is a misperception that open access journals are inferior to paid subscription journals – many open access journals are peer-reviewed and have a high impact factor (for example, *the Association of Learning Technology’s Research in Learning Technology* or the Open University’s *Journal of Interactive Media in Education*). A good strategy is to disseminate your evaluation research findings through a number of channels, using social media to promote your work to a larger audience. However, research findings may only be included in one peer-reviewed publication unless the different papers focus on different aspects of the study.

CHECKLIST ✓

Have you considered ... ?

Your research question(s)	
Your research design (e.g. post, pre-post, longitudinal)	
Appropriate research methods	
The number of participants you need to survey/interview and the sampling strategy you will use for selecting participants	
What type of data you will collect (nominal/ordinal/scale)	
How you will analyse your data	
How you will report and disseminate your findings	

The ethics of evaluation research

For internal evaluation studies, it may not be necessary to submit an ethics review application to the UCL Research Ethics Committee, or gather participants’ informed consent to take part in evaluation research. An example of an evaluation which falls into this category would be an end-of-module student survey. However, for anything more complex, it is considered best practice to adhere to ethical research guidelines. This is especially important if you are thinking about potentially publishing the outcomes of your research, as most academic, peer-reviewed journals will not accept paper submissions where ethical approval for the study has not been granted.

Information about the ethical review process for UCL research projects is available at:

<http://ethics.grad.ucl.ac.uk/>

In addition to completing an ethics form, you will need to complete a data protection form. Contact data-protection@ucl.ac.uk for more information.

Essentially, it is good ethical practice to:

- **Provide study participants with a description of the research project** being conducted, including any funding associated with the project, what methods will be used to gather data from participants, and how you anticipate disseminating the findings. Appendix 1 shows a typical information sheet.
- **Gather evidence of informed consent from participants;** participants may be told that in participating in a study (or using an e-learning resource), they are automatically providing consent for their data to be used in the study. Alternatively, you may wish to develop a consent form that participants have to sign and return to you (either before or at the point of data collection). Appendix 2 shows a typical consent form for an interview or focus group.
- **Reassure participants of confidentiality.** You may choose to conduct an anonymous study; for example, a questionnaire that gets sent to all members of a particular cohort without a space for an identifier (name, email or traceable number such as student number). This may be useful in encouraging greater participation in your study, and potentially more authentic responses; however, it will prevent you from doing pre- and post- comparisons of the views of individual students (unless you set up a 'panel' in an online survey). Alternatively, you may wish to include the participants' names, but 'anonymise' them, i.e. assign them an anonymous identifier such as 'Student01' or 'StudentPre01' and 'StudentPost01' if you want to compare the view and experiences of participants before and after an intervention.
- **Inform participants that their participation in the study is voluntary and that they may withdraw from the study at any time.** Despite the desire to attain a valid response rate for a survey, or to ensure that the members of a focus group are representative according to particular criteria, it is not good practice for participants to feel coerced into participation. The whole point about ethical research is that participants are treated with respect and this includes giving them the choice as to whether or not they want to participate.
- **Reassure individuals that the decision not to participate, or to participate and express a particular view, will have no negative repercussions.** This is particularly important when involving students, in terms of their progression through the course and their treatment in general.
- **Reassure participants that their data will be stored in line with the [1998 Data Protection Act](#).** This means keeping paper-based files in locked cabinets and/or a locked room with access restricted to those involved in conducting the study, or storing electronic data files in a secure, password-protected area of a UCL server. Information may be temporarily stored on portable storage devices such as USB sticks or hard drives; however, the devices must be securely encrypted. One software tool that will allow you to encrypt folders or devices is TrueCrypt (www.truecrypt.org). Without the password, encrypted folders will be inaccessible, and encrypted drives will appear as unformatted.

For more information about the ethics of educational research, see the guidelines produced by Hammersley and Traianou (2012) for the British Educational Research Association.

CHECKLIST ✓

Have you considered ...

Who will have access to the data?	
How/where you will store your data?	
Whether you need ethical approval?	

If you decide ethical approval is necessary, i.e. you intend to publicly disseminate the findings of the evaluation, have you ...

Completed an ethics application?	
Completed a data protection form?	
Developed a consent form for participants to sign?	
Written a Plain Language Statement (information sheet) for participants to accompany the consent form?	

Summary checklist

CHECKLIST ✓

Why are you carrying out the evaluation?

To demonstrate stakeholder satisfaction (acceptability)	
To evidence a higher quality learning outcome	
To demonstrate that the teaching and learning activity/resource is 'fit for purpose' (QA)	
To demonstrate value for money	
Other reason(s)	

Who/what will you be consulting for your evaluation?

Learners/course participants		Module/course leaders	
Educators (teaching fellows/ academics/ teaching assistants)		Teaching administrators	
Educationalists (learning and teaching advisors)		Managers	
Learning technology / e-learning professionals		Employers	
Other professionals e.g. library		Other	

Which of the following levels will you be seeking evidence for? (You can measure more than one.)

REACTION	Attitudes / perceived confidence	
LEARNING	Learning / assessed competence	
BEHAVIOUR	Behaviour change (skills in action)	
RESULT	Demonstrable impact	

Which of the following level of evaluation will you be conducting?

Micro-level (proof of concept)	
Meso-level (wider implementation)	
Macro-level (quality assurance and enhancement)	

Which of the following types of interventions will you be evaluating?

Technology-free educational designs	
Technology-enhanced learning (TEL)	
Technology-enhanced physical learning spaces (TEPLS)	

What approach will you take in your evaluation?

Quantitative (generalisable)	
Qualitative (insightful)	
Mixed methods (quantitative and qualitative)	

Have you considered ... ?

Your research question(s)	
Your research design (e.g. post, pre-post, longitudinal)	
Appropriate research methods	
The number of participants you need to survey/interview and the sampling strategy you will use for selecting participants	
What type of data you will collect (nominal/ordinal/scale)	
How you will analyse your data?	
How you will report and disseminate your findings?	

Have you considered ...

Who will have access to the data?	
How/where you will store your data?	
Whether you need ethical approval?	

If you decide ethical approval is necessary, i.e. you intend to publicly disseminate the findings of the evaluation, have you ...

Completed an ethics application?	
Completed a data protection form?	
Developed a consent form for participants to sign?	
Written a Plain Language Statement (information sheet) for participants to accompany the consent form?	

UCL e-learning case studies

ISD e-learning student survey

This biennial survey has a quality assurance/enhancement purpose and seeks to gather students' feedback on core services provided by the Information Services Division (ISD). The 2013 survey revealed that mobile device ownership had increased since the previous survey, requiring core services to be mobile-friendly. Core services were generally rated highly in terms of importance and satisfaction, through some services were rated as having high importance but lower satisfaction, which helped to identify priorities for future development. One example was the Lecturecast lecture recording system. Additional data showed that about a third of students had watched lecture recordings, and evidenced student demand for increasing use of lecture capture. In terms of how technology is being used to support their learning, there was evidence that staff were using e-learning for more than administrative or supplementary purposes, and in fact, were using e-learning in an enhanced, integrated or fully-online capacity. The top three services that students named included the Wi-Fi service, support services, and Moodle/online learning resources. Their least favourite experiences included there not being enough available computers, too few printers or printers not working, and unreliable Wi-Fi in certain spaces around campus. These negative experiences have been subsequently addressed through the implementation of a laptop loan scheme, installation of multipurpose printing/scanning machines and boosting of Wi-Fi around the institution.

LectureTools pilot evaluation

A low-risk [pilot evaluation study of LectureTools](#) was carried out with Jane Britton (Management Science and Innovation), to determine whether LectureTools would promote interactivity in the classroom for a professional course with a small group of 17 students. The research methods used included a survey administered using clickers in the classroom, followed by a focus group discussion with students. Two staff (Jane and Matt Whyndham) were also interviewed. The pilot evaluation revealed some technical issues with the platform, particularly on mobile devices. Students appreciated the interactivity but recognised that the presence of laptops in the classroom created an opportunity for distraction. Staff felt that students' interaction with them was mediated by the technology. On one hand, this was good as it encouraged all students to participate. On the other, the use of LectureTools as a medium for communication substantially changed the nature of communication in the classroom.

The tool has subsequently been observed in larger classes. In Jane Burns' Marketing courses (with up to ~60 students), it was observed that the system significantly slowed due to limited wifi capability in the classroom. It was also used in Chris Blackman's first year Chemistry course with about 100 students. Participation by students was minor; partly because time was not invested at the start of the teaching session, getting students logged on and feeling comfortable with using the system. Secondly, the tool use was stressed as optional. It's therefore important to consider technology use in different teaching contexts, as institutional factors are likely to affect the outcome of the evaluation, whether these relate to IT infrastructure such as wifi capability, or the pedagogical use of the technology as determined by the educator.

@inDUCKtion evaluation

The [@inDUCKtion project](#) was an initiative at the start of the 2013-14 session to inject 'fun' into what otherwise can be considered to be an overwhelming induction experience for new students, led by Janina Dewitz and Mira Vogel from ELE, assisted by Paul Walker from CALT. Students were encouraged to use technology to help familiarise themselves with the campus and different services available, as well as improving their digital literacy. The design of the intervention included a virtual tour of the campus facilitated by scanning QR codes and visiting clues and information on a MyPortfolio site. Another aspect was a photo challenge which encouraged students to take photos of a plastic duck in various locations, and tweet these to @inDUCKtion duck or email them so all photos and comments could be curated using Storify.

The evaluation was extensive and involved going into public UCL spaces to interview students, as well as using Twitonomy to look at the @inDUCKtion tweets, mentions and replies. Other quantitative data included a log of QR code scans. A social network analysis was also attempted, with limited success due to lack of familiarity with the platforms and realising that the Twitter API search function would only go back so far in time. Key staff members in ELE, CALT and Student Advisory and Event Services.

Although a relatively small number of students engaged with the initiative, the project was less successful than anticipated due, again, to institutional factors as well as personal factors. The main institutional limitation was that the project took place within an already overwhelming induction period. The main personal factor was that, surprisingly, a number of students did not have the technology or did not know how to scan QR codes. One of the main recommendations arising from the evaluation was to secure buy-in from student representatives and integrate any future initiatives into departmental induction activities, to promote institutional endorsement.

Blikbook evaluation

Blikbook is a Question and Answer tool that was first adopted by Richard Pettinger, Principal Teaching Fellow in Management Science and Innovation (MS&I). Given the apparent success of Blikbook within this department, the developers of the tool approached UCL with a proposal to encourage UCL to adopt Blikbook as a core supported service across all departments. To determine whether this was an appropriate strategy, an evaluation was carried out comprising a feature comparison with comparable platforms, interviews with UCL educators in four different departments who had used Blikbook in their teaching, a survey targeted at their students, and an examination of the system log data to assess the extent to which it had been used.

Blikbook offered certain advantages over other tools such as its modern appearance, tie-in deal with a publishing company to distribute e-books, as well a useful search function, and the ability to post questions anonymously. However, at the time of the evaluation, the publishing facility was not yet available, users were unable to post PDF documents as attachments, and Blikbook was not being used to any significant extent in the other three departments. It did not perform as well as other platforms in the feature comparison or in the student survey, and educators also expressed mixed feelings about the ability to track individual users through learning analytics, one of the key features of Blikbook. As a result, it was decided that it was not appropriate at this time to adopt the platform as a centrally-supported core resource for the whole of UCL. However, the staff using this platform were acknowledged in the report for their innovative practice, and encouraged to continue to

explore new e-learning solutions in partnership with ELE in the context of institutional small-grant funded e-learning projects.

Making History group project (and MyPortfolio) evaluation

In 2013-14, the Department of History at UCL implemented a new first year curriculum. The first term comprised two core modules, including one called Making History. This module was designed to engage students in research-based learning from Day 1 of their undergraduate degree, as a basis for developing historical skills, digital literacies and other employability skills. Students were required to create an interactive home page in MyPortfolio, linking to essays and a weekly log of group activities. The evaluation comprised a student survey and an interview of educators including course leaders and postgraduate tutors.

The evaluation revealed that, in general, the module was a success. Tutors had attempted to 'shake students out of A-level thinking' and as a result, students had engaged in research-based learning and developed a range of employability skills including digital literacies and history-specific skills. Experiences of MyPortfolio (Mahara) were generally straightforward although students had some difficulties embedding external files and struggled to design their project home page in the way that they wanted to. Although the groupwork experience was generally positive, there was evidence of some dysfunctional teamworking including social loafing. Tutor commented that the groupworking accounts in the teamworking essay and weekly logs were essentially 'sanitised'. In addition, it was clear that students had worked to their own strengths, making it possible that not all students had developed the desired skills. These negative findings were partly attributable to the assessment strategy, which awarded group members a single identical mark for the whole endeavour. Recommendations for the programme next session are to include an individual mark for each student awarded through peer assessment, as well as an introduction to the principles of teamwork so that students are better prepared to handle teamwork challenges and do not feel the need to hide teamworking difficulties.

Glossary

Coding – in the context of a survey, ‘coding’ refers to assigning a numerical value to a particular response, for the purposes of data analysis in a statistical package such as SPSS. An example might be Yes=1, No=0 or 1=Scotland, 2=England, 3=Ireland and 4=Wales. In qualitative analysis, coding refers to labelling a segment of text for the purpose of aggregation and comparison. For example, the sentence “When I review my lectures before the exam, I also produce a concept-map of the main concepts and the relationships between them” might be coded as ‘Revision strategies’.

Delphi process – The Delphi process is a way of harnessing expert knowledge. Usually, panel members are asked to rate a series or items on a survey according to importance, or the survey may seek qualitative data. The collective data is then presented back to the panel participants, who revise their individual ratings or answers based on their consideration of other panel members’ collective responses. This process may undergo two or more rounds until a consensus has been achieved. The predictions in the NMC Horizon Report is based on this process.

Focus group – A semi-structured discussion with a group of 6-8 individuals which capitalises on the interaction between the group members.

Formative evaluation – At this stage, evaluation data is collected as a course/module/initiative is ongoing, and refinements and corrective actions may be undertaken in time to benefit the current cohort(s) of students/staff/other stakeholders.

Group interview – A group interview is similar to a focus group but does not capitalise on the opportunity for participants to engage in a discussion with each other, negotiating their expectations and experiences. In a group interview, a question is asked and all participants have the opportunity to answer the question in turn.

Interview – The interview is typically carried out one-to-one between a single research participant and the researcher. Interviews may be structured, semi-structured or unstructured. A structured interview has a rigid structure of set questions, from which there are no deviations. An unstructured conversation is free-flowing and conversational in style. A semi-structured approach is best, where the researcher has specified a number of prompts in advance, but still allows the participant some freedom to introduce topics of interest to them.

Nominal Group Technique – The Nominal Group Technique is a consensus method that can work well in the small group setting. Participants may be invited to write down their ideas on individual post-it notes, which are then presented for the rest of the group to see. The participants add further ideas and are asked by the session facilitator to organise the post-it notes and agree on priority areas. The facilitator then elicits further discussion about these areas.

Non-parametric statistics – These are used to look for statistical differences or associations between categorical variables (data resulting from dichotomous, multiple choice or Likert scale questions). Commonly used tests include the Mann-Whitney U test and the Kruskal Wallis ANOVA to compare two or more groups respectively, and the Spearman’s Rho correlation test.

NVivo – NVivo is a qualitative analysis program available for download via the UCL software database (www.ucl.ac.uk/isd/common/software). NVivo will not do anything that you cannot do on

paper, in terms of coding segments of text and identifying recurring themes in the data, but it makes the process of storing and querying the data much easier. Its use is recommended particularly for large data sets.

Parametric statistics – Parametric statistics are typically conducted with data that is continuous and normally distributed i.e. when the data is represented as a histogram, it appears as a ‘bell shaped curve’. When the data is categorical or not normally distributed, then non-parametric tests should be used.

Qualitative research – Qualitative research approaches are grounded in the interpretivist paradigm, and typically gather non-numerical data such as open text survey responses, interview and focus group transcripts. Quantitative data are subjected to thematic analysis – though this is a broad term which encompasses of a range of analytic techniques, some of which make particular assumptions about the nature of knowledge and how it is constructed or the social context in which the research is being conducted.

Quantitative research – Quantitative research approaches are grounded in the positivist (hard sciences) paradigm, and typically gather numerical data which may be subjected to statistical analyses. The most common quantitative research method is the questionnaire, but system usage logs would also constitute quantitative data.

Response rate – It is important to report the response rates to surveys to indicate how representative of the population the results are. For online surveys, a response rate of between 30-40% is typical. You will need to acknowledge any potential biases in describing the limitations of your study, and how you took steps to try to eliminate these.

Sampling – Typically it will not be possible to survey, interview or observe all members of a population, whether these are students, educators or other stakeholders, thus a decision will need to be made as to how many participants, and what type of participants, to sample. For example, will the sample constitute a group of similar individuals (to gather information from this specific group) or a diverse group, to provide a range of opinions/experiences?

Summative evaluation – At this stage, evaluation data is collected at the end of a course/module/initiative. It will not be possible to undertake refinements and corrective actions in time to benefit the current cohort(s) of students/staff/other stakeholders, but future stakeholders will benefit from adjustments made. Evaluation carried out at this stage typically has a quality assessment role, ensuring that resources/facilities are ‘fit for purpose’ as benchmarked against existing quality indicators.

References

- Ala-Mutka, K. (2009). "Review of Learning in ICT-enabled Networks and Communities." Retrieved 23 July, 2014, from <http://ipts.jrc.ec.europa.eu/publications/pub.cfm?id=2721>.
- Cohen, L., L. Manion and K. Morrison (2007). Research Methods in Education. London, Routledge.
- Field, A. P. (2009). Discovering statistics using SPSS: (and sex and drugs and rock 'n' roll). London, SAGE.
- Hammersley, M. and A. Traianou. (2012). "Ethics and Educational Research, British Educational Research Association on-line resource." Retrieved 24 July, 2014, from <http://www.bera.ac.uk/wp-content/uploads/2014/03/Ethics-and-Educational-Research.pdf>.
- Harvey, J. (1998). "LTDI Evaluation Cookbook." Retrieved 15 July, 2014, from <http://www.icbl.hw.ac.uk/ltdi/cookbook/>.
- Kitzinger, J. (1995). "Qualitative research: introducing focus groups." British Medical Journal **311**(7000): 299-302.
- Moore, D. A., D. J. Klingborg, J. S. Brenner and A. A. Gotz (2002). "Using focus groups for continuing veterinary medical education needs assessment and program planning." Journal of Veterinary Medical Education **29**(2): 101-104.
- Morrison, K. (2001). "Randomised controlled trials for evidence-based education: some problems in judging 'what works'." Evaluation & Research in Education **15**(2): 69-83.
- Nulty, D. D. (2008). "The adequacy of response rates to online and paper surveys: what can be done?" Assessment & Evaluation in Higher Education **33**(3): 301-314.
- Pallant, J. (2010). SPSS Survival Manual: A Step by Step Guide to Data Analysis Using SPSS. Maidenhead, Open University Press.
- Pearshouse, I., B. Bligh, E. Brown, S. Lewthwaite, R. Graber, E. Hartnell-Young and M. Sharples. (2009). "A study of effective evaluation models and practices for technology supported physical learning spaces (JELS)." Retrieved 23 July, 2014, from <http://oro.open.ac.uk/29996/2/B0DAC2BE.pdf>.
- Robson, C. (2011). Real World Research: A Resource for Users of Social Research Methods in Applied Settings. Oxford, Wiley-Blackwell.
- Roberts, S. and M. Weaver (2006). "Spaces for learners and learning: evaluating the impact of technology-rich learning spaces." New Review of Academic Librarianship **12**(2): 95-107.
- Vagias, W. M. (2006). "Likert-Type Scale Response Indicators." Retrieved 13 July, 2014, from <http://www.clemson.edu/centers-institutes/tourism/documents/sample-scales.pdf>.
- Varga-Atkins, T., N. Bunyan, R. Fewtrell and J. McIsaac. (2011). "The Nominal Group Technique – a Practical Guide for Facilitators." Retrieved 15 July, 2014, from http://www.liv.ac.uk/cllserver/resource-database-uploads/pdf/guide_for_ELESIG_v1.pdf.

Wilson, G. and M. Randall. (2012). "The implementation and evaluation of a new learning space: a pilot study." Research in Learning Technology Retrieved 24 July, 2014, from <http://www.researchinlearningtechnology.net/index.php/rlt/article/view/14431>.

Additional reading

Jisc. (2007). "Methods for Evaluating the Learner Experience of E-Learning." Retrieved 15 July, 2014, from http://www.jisc.ac.uk/media/documents/programmes/elearningpedagogy/guide1_evaluation.pdf.

Jisc. (2010). "E-Learning Programme and Project Evaluation." Retrieved 15 July, 2014, from <http://www.jisc.ac.uk/whatwedo/programmes/elearningcapital/evaluation.aspx>.

Phillips, R., C. McNaught and G. Kennedy (2012). Evaluating E-Learning: Guiding Research and Practice. London, Routledge.

Ruhe, V. and B. D. Zumbo (2009). Evaluation in Distance Education and E-Learning: the Unfolding Model. New York; London, Guilford.

Appendix 1: Example information sheet



Information Sheet: Evaluation of the Making History group project

This study has been approved by the UCL Research Ethics Committee (Project ID Number): 4601/003

Dear Student / Postgraduate tutor,

We would like to invite you to participate in a study to evaluate the Making History group project.

This study seeks to gather information from 1st year students and their postgraduate tutors, to evaluate the benefits and limitations of the Making History group project, and to identify what additional support might be required for this project to run effectively in future.

Participation in the study is entirely voluntary. Refusal to participate will involve no penalty or loss of benefits to which participants are otherwise entitled. Participants may discontinue their involvement in the study at any time without penalty or loss of benefits.

The main methods used will be an interim and final student focus group (during and after the project), individual interviews with postgraduate tutors, a survey disseminated to all students, and content analysis of the weekly logs and the resulting presentations in MyPortfolio.

A scheduled proforma will be provided to participants in advance of the focus groups/interviews, so that they may think about the questions in advance. The focus groups/interviews will be audio-recorded for the purposes of clarifying information provided, and including representative quotes in the final report. The audio files will only be accessible to the researchers named below and will be deleted after the project ends, by the December 1st 2014.

We anticipate that participants will benefit from involvement in the study in that learners will be able to reflect on what they have learned, and how they are developing as historians, and tutors will be able to reflect on their teaching/supervisory role in terms of what is working well and what might be refined, with support.

The information provided will be stored securely in accordance with the Data Protection Act 1998. Only the main researchers (Drs Mira Vogel and Vicki Dale) will be able to identify the notes/quotes back to individual participants. The other researchers will have access to anonymised results, and the final study outcomes will be reported anonymously, unless participants have given permission to be named.

While this study will conclude by 1st December 2014, at which time audio recordings and identifiable transcripts will be deleted/shredded, we would like to maintain the raw, unidentifiable data in a password-protected area of a UCL server for additional or subsequent research and to inform the future development of the Making History group project and similar modules.

If participants have questions about the study, they may contact us at any time.

Finally, we would like to thank you in advance for your participation in the study.

Dr Mira Vogel, E-Learning Facilitator (SLASH) and Dr Vicki H.M. Dale, E-Learning Evaluation Specialist

E-Learning Environments, ISD, Gower Street, London WC1E 6BT

Tel: +44 20 354 95197, Email: m.vogel@ucl.ac.uk

Please discuss the information above with others if you wish or ask us if there is anything that is not clear or if you would like more information.

It is up to you to decide whether to take part or not; choosing not to take part will not disadvantage you in any way. If you do decide to take part you are still free to withdraw at any time and without giving a reason.

All data will be collected and stored in accordance with the Data Protection Act 1998.

Appendix 2: Example consent form

Informed Consent Form

Please complete this form after you have read the Information Sheet and/or listened to an explanation about the research.

Title of Project: **Evaluation of the Making History group project**

This study has been approved by the UCL Research Ethics Committee (Project ID Number): 4601/003

Thank you for your interest in taking part in this research. Before you agree to take part, the person organising the research must explain the project to you.

If you have any questions arising from the Information Sheet or explanation already given to you, please ask the researcher before you to decide whether to join in. You will be given a copy of this Consent Form to keep and refer to at any time.

Participant's Statement

I

- have read the notes written above and the Information Sheet, and understand what the study involves.
- understand that my participation is voluntary, and that if I decide at any time that I no longer wish to take part in this project, I can notify the researchers involved and withdraw immediately.
- understand that my participation will be audio-recorded and I consent to use of this material as part of the project.
- consent to the processing of my personal information for the purposes of this research study.
- understand that such information will be treated as strictly confidential and handled in accordance with the provisions of the Data Protection Act 1998.
- (as a student) understand that the evaluation is being carried out by e-learning staff with no influence or responsibility for my assessment results. Any focus group data provided will be anonymised and any survey data will be aggregated before being communicated to history staff *after* the assessment process.
- agree that the research project named above has been explained to me to my satisfaction and I agree to take part in this study.
- understand that the information I have submitted will be published as a report and I will be sent a copy. Confidentiality and anonymity will be maintained and it will not be possible to identify me from any publications.
- agree that my non-personal research data may be used by others for future research. I am assured that the confidentiality of my personal data will be upheld through the removal of identifiers.

Name (please print): _____

Signature/Initials: _____

Date: _____

Creative commons licensing



UCL E-Learning Evaluation Toolkit by [Vicki H.M. Dale](#) is licensed under a [Creative Commons Attribution-NonCommercial- 4.0 International License](#).