



UNIVERSITY OF  
**Southampton**  
School of Electronics  
and Computer Science



The Open University

# **Synthesis Report on Assessment and Feedback with Technology Enhancement**

**July 2011**

**Prepared by  
Lester Gilbert<sup>1</sup>, Denise Whitelock<sup>2</sup> and Veronica Gale<sup>1</sup>**

<sup>1</sup> Learning Societies Lab in the School of Electronics and Computing Sciences,  
University of Southampton

<sup>2</sup> Institute of Educational Technology, The Open University, UK

## **Executive Summary**

This report has been commissioned by The Higher Education Academy (The Academy) in January 2010. The purpose of the project was to:

- Consult the academic community about which references on assessment and feedback with technology enhancement are most useful to practitioners;
- Synthesise the main points from these references;
- Provide signposts for readers to the original documents for further study.

The target readers for this report are:

- Academics who are using technology enhancement for assessment and feedback;
- Learning technologists working with academic colleagues;
- Managers of academic departments.

The report addresses the main questions this audience are likely to have about assessment and feedback with technology enhancement, i.e., What can I do? How can I do it? What evidence can I use to convince stakeholders?

The report is structured around the steps involved in the creation and use of technology-enhanced assessment and feedback. The majority of the references concern the first two stages in the process: deciding to use technology to achieve a specific goal, and the design of the assessment and feedback. Much of the effective practice described in the report has been developed for a specific technological approach such as using podcasts, online discussion forums or multiple-choice questions. However the design, project management and quality assurance principles identified in this report may be applied to other types of technology enhancement as well.

## **Findings**

The brief was to identify which publications practitioners find particularly useful amongst the many materials and papers available on technological support and enhancement for assessment and feedback, and then to focus on those references that were evidence-based, widely available and peer-reviewed. Of the 142 references which were recommended, the report reviews the 124 which were accessible, including journal articles, reports, books, and websites.

We asked practitioners for references that they would recommend to other practitioners. We explained that we were particularly interested in 'evidence-based' publications. We anticipated the 'evidence' that would be provided would be high-quality statistical analyses showing that actual change had taken place in either learning, performance, or at an organisational level. However, while all reviewed references were from reputable sources and the majority were peer-reviewed (67.7%), only a minority provided quantitative data (28.2%), of which relatively few provided the experimental designs or statistical analyses (18.5%) needed to show that technology had provided a measurable enhancement. Most references focused on the reaction of students and teachers to the use of technology for

assessment and feedback. Although it may be ideal to have high-quality evidence before implementing a new assessment approach, in the absence of this level of support, the insights and advice of reputable authors are valued by other practitioners and their messages have impact.

The report provides a detailed introduction to and summary of references useful to practitioners on technology-enhanced assessment applicable to Higher (and Further) Education in a UK context.

## **Key messages**

The practitioners we consulted recommended articles that demonstrate the effectiveness of technology-enhanced assessment, challenge assumptions about the use of technology-based methods or give clear support for specific learning designs.

The review of the surveys and case studies that were considered valuable showed that technology can enable Higher Education Institutions (HEIs) to deliver the characteristics of assessment that make it effective, such as frequent, on-demand formative tests with tailored feedback. Use of technology can enhance teaching by making effective learning designs such as these possible. Without the use of technology, constraints such as time, cost, student numbers, and geographical or temporal distribution would make these learning designs impractical to deliver.

Here are just some of the messages that are supported by evidence in the recommended literature:

- Assessment for learning shows an effect size of between 0.34 and 0.46, see page 8;
- Tutors can use technology-enhanced methods to implement effective learning designs that would not otherwise be possible because of factors such as time constraints, student numbers and geographical or temporal distribution, see page 9;
- Effective regular, online testing can encourage student learning and improve their performance in tests, see page 10;
- Student retention and inclusion can be increased by using technology-enhanced methods. Exam anxiety can also be reduced, see page 17;
- Using technology-based methods does not disadvantage women or older students, see page 21;
- Automated marking can be more reliable than human markers and there is no medium effect between paper and computerized exams, see pages 16, 22 and 23;
- The success of assessment and feedback with technology enhancement lies with the pedagogy rather than the technology itself; technology is an enabler, see page 26;
- Technology-enhanced assessment is not restricted to simple questions and clear-cut right and wrong answers, much more sophisticated questions are being used as well, see page 28;
- Modern technology can be matched to the learning characteristics of the contemporary learner, see page 28;
- The design of appropriate and constructive feedback plays a vital role in the success of assessment, especially assessment for learning. The literature offers detailed guidance on designing effective feedback such as conditions, research-backed

principles and a typology, as well as specific advice for the design of audio feedback and peer assessment, see page 33;

- What the characteristics of useful technologies to use for assessment and feedback are, see page 40;
- Taking a team approach to the creation of technology-enhanced assessment and feedback is valuable because successful implementation requires skills in the application of technology and how to use to the technology itself as well as learning and the subject content, see page 42;
- Staff development and support are vital when introducing and developing assessment and feedback with technology enhancement, see page 43;
- Testing the assessment and feedback to ensure it is reliable and valid and piloting it with people who are similar to or understand the targeted students are important stages in the development process. A good reporting system can help academics see and analyse the results (including student answers) and will help refine the assessment and feedback, see page 45;
- It is important to prepare students to take the assessments that use technology enhancement by practising with similar levels of assessment using the same equipment and methods. This is similar to being able to practise on past papers, see page 48;
- The reports generated by many technology-enhanced assessment systems are very helpful in checking the reliability and validity of each test item and the test as a whole, see page 49.

## **Conclusions**

The references that were recommended to us are clearly having an impact on current practice and are found valuable. We would welcome more high-quality statistical studies that offer evidence to support the lessons that practitioners have learned from experience. It would also be interesting to do further work on why these particular articles have an impact on practitioners.

## Contents

|  |    |
|--|----|
| <b>I. Introduction</b> .....   | 1  |
| A. Background.....   | 1  |
| B. Acknowledgements.....   | 1  |
| C. Methodology .....   | 1  |
| D. Assessment and feedback context .....   | 2  |
| E. Technology enhancement .....  | 3  |
| F. Evidence-based literature .....   | 3  |
| G. Report structure.....   | 4  |
| <b>II. Key messages in recommended literature</b> .....                              | 7  |
| 1. Using technology to meet the goals set for assessment and feedback .....          | 8  |
| 1.1 Introduction.....  | 8  |
| 1.2 Meeting the goal of improving student performance .....                          | 10 |
| 1.3 Meeting the goal of encouraging self-regulation .....                            | 12 |
| 1.4 Meeting the goal of encouraging student reflection .....                         | 13 |
| 1.5 Meeting the goal of facilitating student collaboration and peer assessment ..... | 14 |
| 1.6 Meeting the goal of making assessments more authentic to the task .....          | 14 |
| 1.7 Meeting the goal of saving time, effort and costs .....                          | 15 |
| 1.8 Meeting the goal of offering more flexible assessment arrangements .....         | 17 |
| 1.9 Meeting the goals of increasing student retention and inclusion .....            | 17 |
| 1.10 Case studies of technology meeting stated goals.....                            | 18 |
| 1.11 What has limited the take-up of technology to meet goals? .....                 | 20 |
| 1.12 Summary of using technology to meet learning goals .....                        | 23 |
| 2. Designing assessment and feedback with technology enhancement.....                | 25 |
| 2.1 Introduction.....  | 25 |
| 2.2 Design should be driven by pedagogy rather than technology .....                 | 26 |
| 2.3 Designing assessment questions and feedback for technology enhancement.....      | 27 |
| 2.4 Designing assessments for the contemporary learner .....                         | 28 |
| 2.5 Designing the use of e-portfolios .....  | 29 |
| 2.6 Designing technology enhancement for group work and peer assessment.....         | 31 |
| 2.7 Designing feedback for a technology-enhanced environment.....                    | 32 |
| 2.8 Summary for designing technology-enhanced assessment questions and feedback..    | 38 |
| 3. Creating the assessment and feedback using technology .....                       | 40 |
| 3.1 Introduction.....  | 40 |
| 3.2 Characteristics of useful technologies .....                                     | 40 |
| 3.3 Using a team approach to the creation of assessment and feedback .....           | 42 |

|   |           |
|---|-----------|
| 3.4 Staff development and support needed when using technology .....              | 43        |
| 3.5 Summary for creating assessment and feedback with technology enhancement..... | 44        |
| 4. Testing .....  | 45        |
| 5. Piloting .....   | 45        |
| 6. Delivery .....   | 46        |
| 7. Capturing Student Responses .....  | 48        |
| 8. Feedback to Students.....  | 48        |
| 9. Evaluation .....   | 48        |
| 9.1 Measuring validity and reliability .....                                      | 48        |
| 9.2 Evaluating reports generated by the assessment technology .....               | 49        |
| 9.3 The importance of regular reviews .....                                       | 49        |
| 9.4 Summary of key messages on the evaluation stage.....                          | 50        |
| 10 Feedback to Stakeholders.....  | 50        |
| 10.1 Introduction.....  | 50        |
| 10.2 Addressing the different interests of different stakeholder groups.....      | 50        |
| 10.3 Ensuring accessibility.....  | 51        |
| 10.4 Ensuring high quality.....   | 51        |
| 10.5 Summary of key messages on feedback to stakeholders.....                     | 52        |
| <b>III. Conclusions .....</b>   | <b>53</b> |
| A. Limitations .....  | 53        |
| B. Observations.....  | 53        |
| C. Comments .....   | 53        |
| D. Further work.....  | 54        |
| E. Key messages .....   | 54        |
| <b>IV. Appendices .....</b>   | <b>56</b> |
| A. Survey questions .....   | 56        |
| B. Glossary.....  | 56        |
| C. Lists of principles and conditions .....                                       | 57        |
| D. Academy's Seminar Series links.....  | 59        |
| E. Zakrewski and Steven (2003): interests of different stakeholder groups .....   | 59        |
| F. Cost-benefit analysis for e-learning .....                                     | 60        |
| G. Kirkpatrick levels of evaluation.....  | 61        |
| H. Evidence category 1a references.....   | 61        |
| <b>V. References .....</b>  | <b>63</b> |

## **Tables**

|   |    |
|---|----|
| Table 1: Categories of evidence used in this report                 | 4  |
| Table 2: Number of references recommended in each evidence category | 4  |
| Table 3: Case studies of using technology to meet specific goals    | 18 |
| Table 4: Typology of formative assessment                           | 33 |
| Table 5: Case studies of designing technology-enhanced feedback     | 36 |
| Table 6: Examples of software tools developed by HEIs               | 40 |

## **Figures**

|   |    |
|---|----|
| Figure 1: Cycles of e-assessment from Whitelock (2006a) e-Assessment: Case Studies of Effective and Innovative Practice | 5  |
| Figure 2: Report structure -- Stages in the process of creating and using technology-enhanced methods                   | 6  |
| Figure 3: The Instructional Transaction   | 25 |
| Figure 4: Dialogic feedback cycle, Beaumont et al (2008)  | 32 |
| Figure 5: Example report from the PASS-IT project   | 41 |
| Figure 6: Quality Assurance 'Catherine Wheel' from Zakrewski and Steven (2003)  | 52 |

## **I. Introduction**

### **A. Background**

This report was commissioned by The Higher Education Academy (The Academy) in January 2010. It is one of three synthesis projects that The Academy has commissioned. The others are on Professional Learning and Teaching Practice, and Employability and Employee Learning. The aims of the synthesis initiative are to support the sector in its use of technology to enhance learning and teaching by providing a comprehensive and useful synthesis of evidence-based practice.

The purpose of the SRAFTE project was to:

- Consult the academic community about which references on assessment and feedback with technology enhancement are most useful to practitioners;
- Prioritise evidence-based references i.e., those that are peer reviewed and have data to support their practice;
- Synthesise the main points from these references;
- Provide signposts for readers to locate the original documents for further study.

The target readers for this report are:

- Academics using technology enhancement for assessment and feedback;
- Learning technologists working with academic colleagues;
- Managers of academic departments.

### **B. Acknowledgements**

We wish to thank our advisers on the project:

- Helen Ashton, Heriot-Watt University;
- Don MacKenzie, Derby University;
- Gary Wills, Southampton University.

We also wish to thank the following for recommending references to be included:

- Sally Jordan, The Open University, UK;
- David Walker, Dundee University;
- Bill Warburton, Southampton University;

and the participants at the Academy's Assessment and Feedback Seminar Series who completed our questionnaire.

### **C. Methodology**

The literature on technology-enhanced assessment is large and varied. In order to focus on references that practitioners find particularly helpful and would recommend to their peers, we consulted the HE community via:

- An Advisors' Panel;



- The related seminar series organised by the Academy;
- A stand at the CAA 2010 conference;
- Email requests to selected HE practitioners.

We used a questionnaire (Appendix A) on paper and online at Survey Monkey to elicit the information from participants at the Academy's seminar series and at the CAA 2010 conference. We also had a cloud on Cloudworks that directed viewers to the survey online.

We received 142 references, including journal articles, reports, books and websites. These were reviewed to identify:

- The technology-enhanced methods discussed (as opposed to assessment and feedback in general);
- The use of technology for assessment and feedback (rather than other aspects of student learning, course administration, or content management);
- The type of evidence that was provided to support their findings or observations;
- How easy the reference was to access (we were unable to source 18 references, particularly those published within a specialist group, a long time ago, or in another country).

After reviewing the literature we decided to include examples of technology-enhanced 'feed forward' as well as feedback.

## **D. Assessment and feedback context**

This report focuses on literature where any one part of the assessment and feedback system is enhanced with technology.

In this system, assessment is driven by the intended learning outcomes and is considered an integral part of the course's learning design (not as an add-on activity at the end of the course). The intended learning outcomes (ILOs) specify the competency to be demonstrated and the subject matter to be covered. Tutors design assessments to test how the learner's performance matches up to the intended learning outcomes and, where given, the feedback needed to reinforce their learning or to help them achieve the ILOs after further study. The learner taking the assessment supplies their answer to the virtual learning environment (VLE) or the other technology that is being used to enhance the assessment and feedback. The teacher then marks the assessment and provides the mark and any other feedback to the learner via the VLE and/or other technology. The learning design for the course or unit is created by the teacher or tutor who may be working alone or with colleagues who specialise in the subject content, the IT skills to create and deliver the assessment and feedback, or the use of technology as a learning tool. Technology may be used to enhance the delivery of the assessment items, the support for the items (including feed forward of advice for students undertaking the tasks), feedback, and marking.

The report raises salient issues around both summative and formative assessment uses of technology enhancement. The report considers that the distinction between these uses usually refers to the extent to which the assessment feedback comprises feedback for learning in its timing, quality, and quantity. Assessment may also be classified as diagnostic, ipsative, or synoptic, and again the report is less concerned with such distinctions than with

the quality of the feedback involved and its use for supporting learning. We have highlighted where the recommended literature identifies practical differences between these various classifications.

## **E. Technology enhancement**

Our brief focused on assessment and feedback with technology enhancement. 'Technology enhancement' was interpreted as using any method that involved a computer or other technology (such as podcasting) in providing advice before the assessment, during the setting of the assessment activity, supporting the assessment activity, capturing student responses, and providing generic or individual feedback on performance. Examples of the technologies discussed in the recommended texts include audio and video feed forward and feedback; e-portfolios; tests delivered, answered, and marked by computers; electronic voting systems; web-enabled group discussions; and peer reviews.

Technology use that was excluded from the project included online submission of written work and computer collection of marks.

'Technology enhancement' suggests that using the technology provides better quality than the alternative, perhaps paper-based materials or a lecture. Although a case can be made that technology does enhance learning and teaching quality in some cases, technology 'enhancement' may also be in terms of cost savings or productivity improvements (Appendix E).

## **F. Evidence-based literature**

We requested examples of evidence-based literature when we consulted the HE community. By 'evidence-based' we meant studies supported by data, including validity and reliability measures, comparisons of learning achieved with and without technology-enhanced features, effect size estimates, and quantified estimates of time saved or effort needed. 'Effect size' is a statistical measure which provides a normalised or standardised measure of the size of an observed effect relative to the underlying error or random variation in the data. It is calculated as the difference between the means of the experimental and control groups, divided by the pooled standard deviation or similar estimate of the data variation. An effect size of 0.3, for example, means that the average score in the experimental group is 0.3 standard deviations above the average score in the control group. Review, 'compilation', or meta-study references were considered 'evidence-informed' and were treated in the same way as evidence-based references.

We developed and refined five categories of grades of evidence (see Table 1). In general, a study meeting a category listed earlier in the table can also match categories listed later.

References to studies were allocated to a particular category based on the information provided in the article or report. When a case study was cited in a 'compilation', meta-study, or review report we tried to find the underlying source but were not always successful. The categories do not imply a hierarchy of value to the reader.

**Table 1**  
*Categories of evidence used in this report*

| Category | Description   |
|----------|---|
| 1a       | Peer reviewed generalizable study providing effect size estimates and which includes (i) some form of control group or treatment (may involve participants acting as their own control, such as before and after), and / or (ii) blind or preferably double-blind protocol. |
| 1b       | Peer reviewed generalizable study providing effect size estimates, or sufficient information to allow estimates of effect size.   |
| 2        | Peer reviewed 'generalizable' study providing quantified evidence (counts, percentages, etc) short of allowing estimates of effect sizes.   |
| 3        | Peer-reviewed study.  |
| 4        | Other reputable study providing guidance.   |

The categories of evidence outlined above may be applied equally to different kinds of study which target a variety of research questions.

**Table 2**  
*Number of references recommended in each evidence category*

| Evidence category | Number of references recommended <sup>(a)</sup> | Cumulative % |
|-------------------|---|--------------|
| 1a                | 15  | 12.1%        |
| 1b                | 8   | 18.5%        |
| 2                 | 12  | 28.2%        |
| 3                 | 49  | 67.7%        |
| 4                 | 40  | 100.00%      |
| Total             | 124   |              |

(a) A total of 142 references were recommended, but 18 were inaccessible.

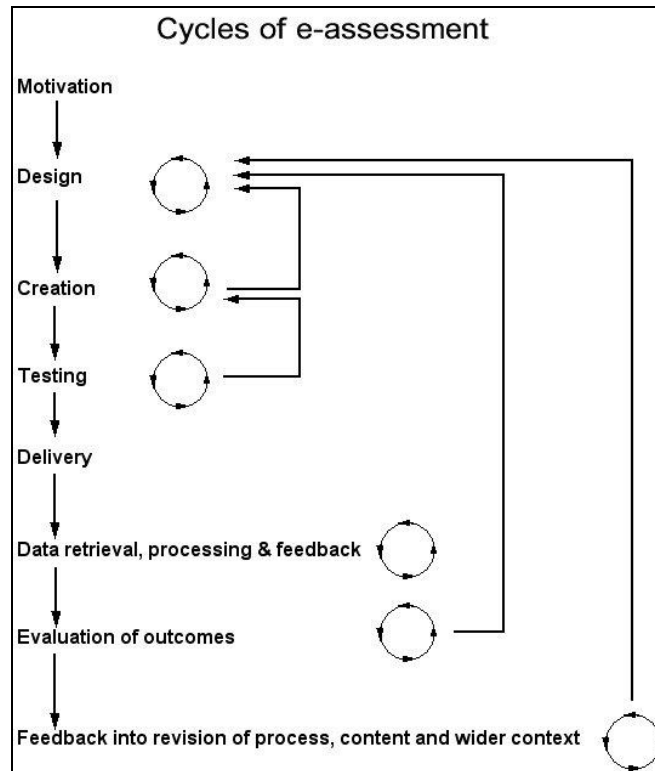
All the recommended references that are classed as evidence category 1a are listed in Appendix H.

For a more detailed discussion of what 'evidence-based' might mean, please see the synthesis report on Professional Learning and Teaching, available from The Higher Education Academy.

## G. Report structure

This is based on the cycles of assessment and feedback shown in Figure 2 developed by Whitelock et al (2006a) in 'e-Assessment: Case Studies of Effective and Innovative Practice', a JISC-funded project conducted by The Open University and the University of Derby (evidence category 3). These are the steps that successful projects followed when creating, refining and improving their technology-enhanced assessment.

Whitelock et al (2006a) highlight the importance of the motivation stage as the 'driver' for the application of technology enhancement against which project outcomes are evaluated. Figure 1 also emphasises the importance of testing/amending the assessment and feedback at each stage in its development, with the outcomes of later stages feeding back into earlier stages for that or subsequent projects.

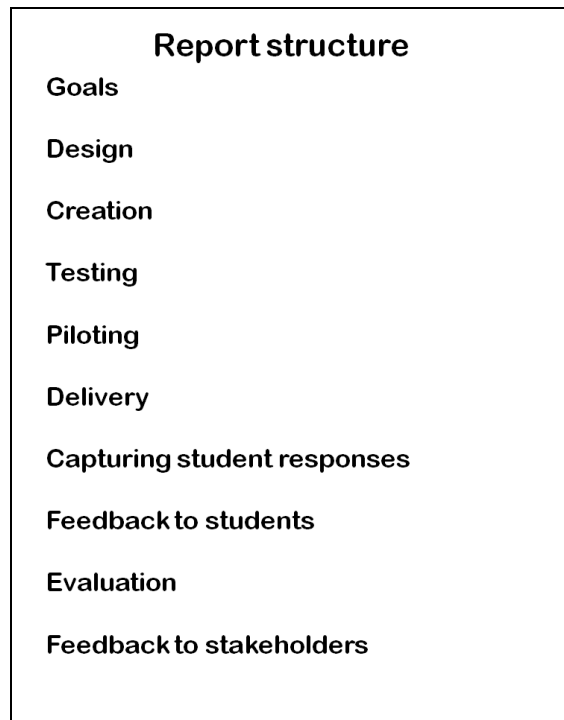


**Figure 1**

*Whitelock (2006a) e-Assessment: Case Studies of Effective and Innovative Practice*

In developing this report's structure, we have modified these stages as follows (Figure 2), by:

- Expanding the first stage to include all the goals that tutors want the course assessment and feedback to meet, such as enhancing student performance, providing the 'human touch' or saving tutor time and effort. This includes the motivation for using technology-enhanced methods while highlighting the need to integrate these tools into the overall course design.
- Dividing the testing stage into (a) ensuring that the technology works, and (b) pilot testing the assessment and feedback against the learning design objectives and intended learning outcomes.
- Separating feedback from data retrieval and processing, as effective feedback is key to the success of many technology-enhanced projects.



**Figure 2**

*Report structure -- Stages in the process of creating and using technology-enhanced methods*

This report's findings are presented under the headings shown in Figure 2, being the relevant steps in the process of creating and using assessment and feedback.

## **II. Key messages in recommended literature**

In this part of the report we summarise the key messages from the literature that practitioners recommended. These messages are presented according to the stage in the development process that successful projects followed when creating, refining and improving their technology-enhanced assessment (Whitelock et al, 2006a; please see previous page).

### **General comments**

The studies that have the greatest impact on practitioners are the ones that demonstrate the effectiveness of technology-enhanced assessment, challenge assumptions about the use of technology-based methods or give clear support for specific learning designs.

Our brief was to find evidence-based studies so we have highlighted those with the strongest data support.

Much of the effective practice described in the report has been developed for a specific technological approach such as using podcasts, online discussion forums or multiple-choice questions. However the design, project management and quality assurance principles can be applied to other types of technology enhancement as well.

Many of the articles that practitioners recommend to their peers are case studies of how an academic team has applied technology to improve assessment and feedback or to address a specific problem. These case studies are valuable, especially to others with similar problems or subject areas where it is clear how to transfer the methods used to other situations. Some case studies, on the other hand, are so specific to a particular situation that it is difficult to see how the lessons learned in that circumstance could apply elsewhere. We have not included such case studies in this report.

There are also useful guidelines (often at a very detailed and generic level, such as how to design an effective multiple-choice question), which project participants recommend for colleagues who are fairly new to the field. These guides are written by experienced practitioners and are often based on applying technology-enhanced methods to many different projects. Understandably, there are no papers giving the details of technology-based projects that failed. However, it is reasonable to assume that the guidelines provided by experienced practitioners include lessons they have learned from projects that did not deliver the results they had expected.

The introduction of technology-enhanced methods often represents a change not only in the conduct of a specific course but also in inter-department working. For example, academics may require the support of the HEI's computer services department and may be offering exam data to administrators in a different format. Further, the policies of the institution may need to be amended to accommodate the technology-enhanced method (such as timing and conduct of exams). These changes need to be managed just as carefully as the technology-enhanced method itself. SRAFTE participants highlighted a number of papers on institutional change. However, this is an area where it is often only possible to transfer generic lessons learned, such as the importance of senior management support and practical arrangements

such as staggered starts for computer-based exams. Less experienced practitioners may find these pointers useful as a checklist of aspects to consider when delivering a new assessment method.

## **1. Using technology to meet the goals set for assessment and feedback**

### **1.1 Introduction**

By 'goals' we mean the objectives that tutors want the assessment and feedback on a particular course to achieve. As well as testing student performance against intended learning outcomes, tutors may have goals such as improving student test scores compared with previous years, encouraging student self-regulation and reflection, facilitating student collaboration and peer assessment, etc. In this section we focus on the evidence for technology-enhanced methods enabling tutors and HEIs to achieve these goals.

We have excluded the references for the impact of particular non-technology-based learning strategies as there is a wealth of literature available which clarifies the relationship between assessment frameworks or strategies and learning and teaching.

Of the references recommended by project participants, 21 discussed assessment and feedback in general, rather than the specifics of technology-enhanced methods. Of these references, some focused on assessment and feedback in higher education (such as Black et al, 2003; Bloxham and Boyd, 2008; Boud and Falchihov, 2007; Irons, 2008; Sotto, 2007), or in higher education and schools (such as Gardner, 2006). All these texts fall into evidence category 4. High quality statistical evidence is available to support how effective formative assessment and feedback has been, especially in schools (Black et al 2003, evidence category 1a), which shows an effect size of between 0.34 and 0.46.

We asked a selection of participants why these references were valuable to them and were told that the principles of effective assessment and feedback hold true for all media, including technology-enhanced ones. The review of the surveys and case studies that had an impact on practitioners showed that technology can enable HEIs to deliver assessment and feedback that meet these principles, such as providing frequent, on-demand formative tests with tailored feedback. Use of technology can enhance teaching by making effective learning designs such as these possible. Without the use of technology, constraints such as time, cost, student numbers, and geographical or temporal distribution would make these learning designs impractical to deliver.

In this section of the report we summarise the characteristics that the cited references identify are needed for effective assessment and feedback. We then highlight examples where technology has been used to provide these characteristics.

Assessment and feedback methods may also be selected to meet resource management and practical goals such as saving time spent on marking at the end of the academic year, and on providing assessments that are authentic to the competence being tested in a cost-effective way. We explore the evidence for using technology to achieve these goals.

We also look at some of the barriers that deter tutors from using technology-enhanced methods and what evidence exists to support or challenge these views.

### **1.1.1 The role of assessment and feedback in achieving student learning goals**

The first stage of creating assessment and feedback, regardless of the medium that will be used, is to specify what they should achieve and how they fit into the overall learning design for a course. For example, assessment and feedback during the course would be specified to help students learn (often called 'formative assessment'), and would be specified at the end of the course to measure the competence students have achieved (often termed 'summative assessment').

Sadler (1998) defines formative assessment as assessment that is specifically intended to generate feedback on performance to improve and accelerate learning. He identifies that feedback must tell the students how they can bridge the gap between their current performance in the assessment task and the desired performance by further learning.

The importance of assessment in gauging and promoting learning has led to the term 'assessment for learning' being coined, where performance data is used to inform the next steps taken by the tutor as well as the learner. The Assessment Reform Group (2002) defines assessment for learning as:

*“the process of seeking and interpreting evidence for use by learners and their teachers to decide where the learners are in their learning, where they need to go and how best to get there”.*

The Assessment Reform Group (2002) also provides 10 “research-based principles” to guide assessment for learning practice. These are reproduced in full in Appendix C. In addition to the points made above, these principles specify that assessment for learning should:

- Focus on how students learn;
- Be sensitive and constructive because any assessment has an emotional impact;
- Take account of learner motivation;
- Develop the learner’s capacity for self-assessment so that they can become reflective and self-managing.

In summary, the following characteristics have been identified for effective assessment and feedback. (See Appendix C for the full lists of principles and conditions.) It:

- Is an activity that tests the students’ current competence of the learning outcomes;
- Provides information about the performance standard that the students will be judged against;
- Helps the student to perform well;
- Provides feedback that explains how the current competence compares with the desired performance at this stage in the course;
- Provides timely feedback so the student can adjust their performance;
- Provides the opportunity to re-take the assessment to improve performance straight away;



- Provides tailored information for the student on the learning action to take to bridge any gap;
- Provides information for the tutor on the students' performance so that subsequent teaching can be tailored to their learning needs;
- Encourages self-assessment, reflection, and active learning;
- Motivates the student to learn;
- Is available at the appropriate time and place.

The next section of the report explores how technology can be used to enhance learning by addressing these different characteristics of effective assessment and feedback.

### **1.1.2 Assessment and feedback as an integral part of the learning model**

In the editorial to the British Journal of Educational Technology special issue *E-assessment: Developing new dialogues for the digital age* Whitelock (2009) argues that technology-enabled assessment needs to be seen as an integral part of the learning model developed for a course and not an add-on activity towards the end. She recommends that academics:

*“Construct a pedagogically driven model for e-assessment that can incorporate e-assessment and e-feedback into a holistic dialogic learning framework, which recognises the importance of students reflecting and taking control of their own learning (p. 199)”.*

In some of the case studies recommended by practitioners, technology has been used to replace the tutor's or another method's role (such as paper-based marks and comments) in an integrated assessment and feedback process. In others, technology has added a dimension to the student experience that was not possible without technology (such as timely provision of detailed feedback). Other case studies have used technology to overcome resources management difficulties to maintain the standard of the students' overall learning experience.

## **1.2 Meeting the goal of improving student performance**

HEIs have used assessment and feedback with technology enhancement to improve the test scores of students. *Exploring Tangible Benefits of e-Learning: Does investment yield interest?* (JISC et al, 2008), a project supported by the Association for Learning Technology (ALT), The Higher Education Academy, and JISC presents 37 case studies of the benefits various HEIs have felt from the use of technology-enhanced approaches. This evidence falls into category 3. The report concludes:

*“There is clear evidence that e-learning offers increased opportunities for formative assessment leading to real benefits in terms of student retention and achievement.”*

It provides evidence from Glamorgan University that student test performance improved and the students' comments showed that they preferred phased, online assessments testing small areas of the syllabus. At Leeds Metropolitan the typical mark improved from “53% to 63%” with students who practiced past e-assessment papers having a mark on average 15% higher than those who did not.

Angus and Watson (2009) also found “robust evidence” (category 1a) for performance gains from regular online testing. They conducted research at the University of New South Wales with over 1,500 students using controls to discount for factors other than the regular tests. They focused on exposure to online testing rather than the results students achieved for the online test. Their main finding was that “exposure to regular (low-mark) online testing significantly improves student learning as measured by the final proctored examination”. In the study, 92% of students agreed that assessment helped their learning. However, the authors quote Ricketts and Wilks (2002) cautioning that the benefits of online testing can be lost if “either automatic marking or online assessment manipulation features are not well thought through,” highlighting that it is how technology-enhanced methods are used that makes the difference, not just the technology itself.

Marriott’s research (2009) suggests possible reasons for the improved performance Angus and Watson discovered as a result of *exposure* to online testing. This research, conducted by the University of Winchester, gathered 103 feedback forms from students on the first year accounting and finance course at the University of Glamorgan (evidence category 2). The survey found that:

- 70% of the students preferred to be “assessed online rather than by traditional methods”;
- over 92% agreed that “phased assessment improves their learning”;
- over 85% agreed that “phased assessment encourages them to work consistently throughout the year”;
- over 78% agreed that “poor performance [in an assessment] motivates them to work harder for the next test (although good performance does not reduce their motivation to work hard)”;
- 91% regarded immediate reporting of results “valuable” and considered “relevant and timely feedback of results as an important aspect in their learning process”. Qualitative comments highlighted that students used the feedback to identify topics they need to study further.

Marriott concluded that phased online assessment “facilitated opportunities for self-assessment and evaluation at an early stage and throughout the year”.

This survey supports the principles identified by Nicol and MacFarlane-Dick (2006) and the assessment conditions identified by Gibbs and Simpson (2004). Both lists fall into evidence category 4 and are reproduced in full in Appendix C.

Nicol (2009) explored how technology-enhanced formative assessment and feedback can be used to develop skills of self-regulated learning in first year students. He analysed two courses that were redesigned with funding from the ‘Re-engineering Assessment Practices (REAP)’ project, (evidence category 1a). Evaluations of these courses provided evidence for “high levels of student satisfaction, performance improvements in final exams when compared with previous years and in workload efficiency gains brought about by the application of technology”. This builds on Nicol’s 2007 paper (evidence category 2) which had similar findings.

Further evidence for performance gains is provided by Hake (1997) who conducted a survey of 6,542 students attending 62 introductory physics courses in high schools, colleges, and universities in the USA. This research is in evidence category 1a. He found that the courses

that used technology enhancement, what he terms “interactive-engagement” (IE), achieved a significantly larger conceptual understanding “gain” when comparing pre- and post-test scores. He concluded that the results:

*“imply that IE strategies enhance problem-solving ability. The conceptual and problem-solving test results strongly suggest that the classroom use of IE methods can increase mechanics-course effectiveness well beyond that obtained in traditional practice.”*

### 1.3 Meeting the goal of encouraging self-regulation

Nicol’s 2009 paper describes student ‘self-regulation’ activities as:

- Students monitoring and regulating their progress towards an assessment goal;
- Students constructing meaning from their studies;
- “Decoding the feedback message, internalising it and using it to make judgements about and modifying their own work” when students receive feedback on their assessments.

Nicol notes:

- Students who do not receive feedback still make progress (Nicol gives the example of students in very large cohorts who receive almost no feedback on their work), so they must be making ongoing judgements about and managing aspects of their own learning.

The Basic Psychology course redesign that Nicol described in this paper used technology-enhanced formative assessment to “encourage regular engagement in study over the timeline of the course, a deeper reading of psychology texts by students and to provide practice and feedback on writing as preparation for the exam. All this was to be achieved without increasing staff workload”. Evaluations of the redesigned assessment (category evidence 1a) showed that these objectives had been met, in particular:

- “Some groups produced written essays online of an exceedingly high standard”;
- There was evidence of “students stimulating and scaffolding each other’s learning” through the use of online discussion boards;
- “Mean exam performance for the course rose from 51.1 to 57.42% [and the] failure rate dropped from 13% in previous years to 5%”.
- “Students were more socially engaged with peers in the lecture class compared to previous years.”

The redesign of formative assessment on the French course produced similar evidence. Students:

- “reported that working regularly on self assessment tests helped them learn more and that they actually worked harder in French than in other subjects”;
- “valued the opportunity for the regular self-assessments and feedback and they particularly valued the speedy return of marks, which helped them direct their own study”;
- “highlighted the flexibility provided by the online tests in relation to the time and place of study”;

- “commented on the instant feedback [of electronic voting systems] and its value in identifying areas of misunderstanding”;

And

- “the progression rate from first year to second year improved from 71.7% to 78% [and] the fail rate dropped from 24% to 4%”.

The evaluation showed that neither course redesign increased staff workload.

Nicol suggests that “balancing structure and self-regulation in the first year is the causal factor” for the learning gains and reduced failure rates found. The REAP project involved redesigning 19 first-year modules with technology-enhanced methods and “11 of these have provided evidence of quantifiable learning benefits.”

A year later, Strang’s (2010) empirical study suggests that international university students use self-regulated process feedback for deep conceptual learning to achieve good grades in an online course. The sample consisted of 1,200 multicultural and domestic students enrolled in an Australian university. He used Hofstede’s (1991) global cultural model to examine cultural differences and the ASSIST model (Entwistle et al, 2004) to look at different approaches to study. The statistical approach adopted was powerful (evidence category 1a) and although feedback was not directly measured, it was indirectly accessed through students’ reported preferences for four different approaches to studying. The findings revealed that students from collectivist and risk-taking cultures were more likely to obtain higher grades because they were willing to use deep or strategic study approaches when evaluating the e-feedback. Deep and strategic study approaches were more effective in processing e-feedback.

## 1.4 Meeting the goal of encouraging student reflection

The *Exploring Tangible Benefits of e-Learning: Does investment yield interest?* (JISC, 2008) report described how Newcastle University use a “flexible e-portfolio tool to help foster a reflective approach to evidencing the achievement of both module-specific and programme learning outcomes”.

Reflection can also be encouraged in final exams by using technology. Williams and Wong’s (2009) comparative study of closed-book, invigilated exams and open-book, open-web exams (evidence category 2) found that the technology-enhanced format of the exam enabled students to “do a quality, referenced piece of work [... with] more time spent on analysis and thinking.”

Evidence from the Inter/National Coalition for Electronic Portfolio Research, (<http://ncepr.org/>), shows that e-portfolios can also be used to encourage student learning and self-reflection. This website provides the results of research which addresses the questions, “What learning is taking place as a function of electronic portfolios?” and “How do we know?” at 10 HEIs. This site falls into evidence category 4 as it covers a range of projects conducted with different methodologies. The latest research is Cohort IV conducted between 2007 and 2010, and is supported by the Higher Education Academy in the UK. For example, the University of Bradford found evidence that using e-portfolios encouraged students to be more proactive and reflective in their studies. Students are quoted as saying:

*"I used PebblePad honestly towards the end of the year before the deadline...I learnt from my mistake...I'm actually using it a lot more this year to begin with."*

And

*"You don't appreciate (what you've learnt) until you actually look back."*

## **1.5 Meeting the goal of facilitating student collaboration and peer assessment**

Barbera (2009) provided evidence in category 1a that students achieved better final results as a result of a combination of peer and co-assessment using a system of linked e-portfolios called 'netfolio'. The study found that adding co-assessment to peer assessment overcame the shortcomings of the latter method and the weaknesses of e-portfolios as an assessment tool. She also noted that the role of the teacher using such technology enhancement changed to one of "observant assessment". Students could tell that the teacher had seen the feedback given by peers (via the message read register). If the teacher made no comments, the students could infer that he or she agreed with the peer assessment provided.

The role of peer assessment in the learning process was explored by Boud (2000) who discussed assessment for lifelong learning and identified the need for "access to learning peers". His study (evidence category 3) found that "The more complex learning is, the less likely that it can be accomplished in isolation from others. We need others for many purposes: To help us check our understanding, see alternative ways of interpreting a task or situation, act as a sounding board for our ideas, provide support when we are feeling low, to identify new sources of information, and to provide other views and judgements."

## **1.6 Meeting the goal of making assessments more authentic to the task**

A key aspect of the intended learning outcomes of a course is the context in which mastery of the competences is expected. This means that to be effective, assessments need to reproduce the context of the skill as closely as possible. *Exploring Tangible Benefits of e-Learning: Does investment yield interest?* (JISC et al, 2008) described assessments at University of Nottingham Medical School where radiographs could be reproduced accurately on screen in full colour. The same level of detail was not possible with black and white printing, and hence the assessment was more authentic to the task and a better test of student competence.

The same publication reported that the University of Glamorgan used technology to create business simulations that were more authentic to the task than paper-based case studies. Both these examples are evidence category 4.

Williams and Wong (2009) emphasised how the use of open-book, open-web final exams enabled the university to assess business skills more authentically as it mirrored the approach they should take to case study problem-solving in real life. This authenticity also "engages students and inculcates deeper and enriched learning".

## 1.7 Meeting the goal of saving time, effort and costs

Whether or not introducing technology-enhanced methods leads to saving time, effort and costs largely depends on the effort needed to create the materials, how reusable they are, and how practical it is to adapt the technology-enhanced materials with minimum changes for later version of the course. The points discussed in this subsection are further developed in Section 3 on Creation.

Joanna Bull (1999) (evidence category 3) highlighted that:

*“Perhaps the most valuable benefit of CAA is the ability to provide focused and timely feedback to students and staff. Feedback can be used to direct future learning, motivate students to investigate other resources and identify students who need additional support. In an environment of rising student numbers, CAA offers the opportunity to give constructive, detailed and consistent feedback to every student, a task which many academics may find difficult to achieve otherwise.”*

Project participants were able to tell us of recent evidence to support these benefits. For example, *Exploring Tangible Benefits of e-Learning: Does investment yield interest?* (JISC et al, 2008) identified that “Probably the most readily quantifiable cost savings were identified in the area of e-assessment where automated marking of exams for large cohorts of students now takes seconds rather than hours.” This statement was supported by evidence (category 2) from:

- Leeds Metropolitan University where e-assessment had been used for 350 students studying Applied Technology and Finance. Using computer-assisted technology meant that summative exam marks were available within 3 hours of the exam ending, saving 120 hours of staff marking time;
- University of Nottingham Medical School found that using e-assessment reduced the marking time for 330 students from 10 hours to “around 2 seconds”.

The JISC report also had an example from Sheffield Hallam University that showed reduced costs in providing feedback to students because using an e-portfolio approach provided regular feedback to students, reducing “the burden of end-loaded marking”.

Boud and Falchikov (2007) pointed out that “There seem to be no studies that map the expenditure of time and effort on different assessment purposes, though it is reasonable to suggest that most time is spent by university staff on marking assignments and examinations that contribute to final grades”. So the possible savings available from using technology-enhanced methods were likely to be significant.

The Leeds Metropolitan example above saved up to £3,000 per cohort in staff time. However, the JISC report did not specify how much staff time was spent on developing the e-assessments. It did specify that, “increased use of reusable learning objects and question banks could reduce initial development time” for future projects. The report noted that a project at the University of Glamorgan using a simulated business environment involved around “three months of academic staff time and six months of multimedia developer time but the reusable design of the game platform means that two further games/case studies have been developed very quickly”.

There is category 1a evidence of technology being used to reduce the marking and feedback load on tutors from Kochakornjarupong and Brna (2010). This study reports the evaluations of McFeSPA systems which supports teaching assistants (normally PhD students) to mark and give feedback on undergraduate computer programming assignments. This meant that the McFeSPA system provided both a “flexible support tool for semi-automated marking of programming assignments” and “scaffolding for improving feedback”.

Some academics are wary of automated marking, fearing this means that only simple, multiple-choice questions can be posed. However, Mitchell et al (2003) described how computers were used to create a progress test in medical knowledge at the University of Dundee requiring free-text responses, as “objective testing is not an acceptable alternative.” The evidence to support their findings was category 1a as they made a direct comparison between the technology-enabled method and the paper-based method used the year before. The authors found that “the staff input to support the [paper-based tests] amounted to approximately 30-man days, *exclusive* of academic preparation of items, standardising content and production of the 30 page exam scripts.” In contrast the computerised marking of the online test took “a few hours. [This] equates to around 30 seconds per student ‘script’ of 270 items.”

Not only was the marking much quicker with technology enhancement, but the marks awarded were more accurate. The authors also found that for the paper-based version of the test “5.8% of the responses had their marks changed by the moderators. In contrast, the computerised version had an error rate of 0.6%.” This efficiency saving is supported by evidence from Beevers et al (1995) whose experience (category 3) also showed that computers are able to “combine speed with accuracy of marking”.

Some may argue that using technology-enhanced approaches merely shifts the effort required from the tutor from the end of the assessment process to the beginning. However, evidence in category 1a is provided by Jordan and Mitchell (2009) of the effort needed to create valid and reliable free text entry questions. They described how experienced question authors were able to use the Open University’s FreeText Author tool after training.

*“The time spent in the initial writing of the question and answer matching varied between a few minutes and several hours, depending on the complexity of the question. Amending the question and answer matching in the light of student responses was even more dependent on the complexity of the question, taking more than a day for some questions.”* The authors note that the questions involved *“tended to be at the more complex end of the short-answer spectrum and that the inclusion of detailed feedback for complex mark schemes added to the time required for development.”*

Another perspective on how technology enhancement can reduce the amount of time tutors spend preparing tailored feedback is offered by Rotheram (2009 evidence category 4). He undertook a project for JISC called *Sounds Good* which asked the question, “Can digital audio be used to give students quicker, better feedback on their work?” He found that:

*“The answer is ‘yes’, in some circumstances. The most favourable conditions seem to be:*

- *The assessor is comfortable with the technology.*

- *The assessor writes or types slowly but records his/her speech quickly.*
- *A substantial amount of feedback is given.*
- *A quick and easy method of delivering the audio file to the student is available.”*

## **1.8 Meeting the goal of offering more flexible assessment arrangements**

Building on the goal of reducing workloads explored in the previous subsection, Mitchell et al (2003) mentioned another student-experience and cost-saving benefit of using technology-enhanced assessment due to the flexibility it offers for timing formative assessments.

“Students who were unavoidably absent on the day of the test (due to illness or work placement) have been able to sit the test with virtually no admin burden for Dundee staff.”

Technology can also offer flexibility when it comes to the location where the assessment takes place. For example, technology can be used to allow students to sit exams in their own home or office. Williams and Wong (2009) undertook comparative research on the efficacy of closed-book, invigilated exams and open-book, open-web (OBOW) exams that can be sat at a location chosen by the student. The results of their study (evidence category 2) found that an “important factor in favour of OBOW was flexibility regarding the location of the exam, as students appreciate the fact that, with an OBOW, they can easily manage their exam schedule to fit better with work and family arrangements. ‘[I] couldn’t really take the course otherwise’, stated one student.” The authors also found that being able to sit the exam in a familiar setting reduced the students’ exam stress and that the OBOW format was “a format relevant to business/professional education”. Importantly, the study also found that the “opportunity” for students to cheat was very similar to that of closed-book, invigilated exams.

There are still more possibilities offered by technology-enhanced methods to create flexible assessment systems. For example, Challis (2005) considered the benefits of adaptive assessment in her paper *Committing to quality learning through adaptive online assessment* (category 3). Although she noted that “the capability is undeniably there”, she found no evidence that it is currently being used in higher education.

## **1.9 Meeting the goals of increasing student retention and inclusion**

### **Retention**

Using technology-enhanced assessment can have more advantages than just enabling regular testing. Leeds Metropolitan (JISC et al, 2008) found that e-assessment led to a “marked improvement in attendance and achievement” with easier recruitment because students are less “apprehensive of taking a Finance subject since they get regular feedback in how they are doing”.

West Suffolk College (Whitelock 2006a) also found that retention and achievement improved when they used technology-enhanced methods. Students on their Chef Apprentice course used digital photographs to demonstrate their practical chef skills (such as food preparation and working in a kitchen). This type of evidence has “broken down barriers experienced by students who have reading and writing difficulties” (Whitelock 2010).



The JISC 2008 report also described how the University of Wolverhampton found attrition rates on their Nursing and Midwifery courses reduced after introducing an e-portfolio system that helped students communicate with one another.

## Inclusion

Being accessible to students who have “a range of medical and other conditions such as recognised physical disabilities and dyslexia through to students who become pregnant during the course” was another benefit of using technology enhancement that is described in JISC et al (2008).

The need to ensure accessibility for all is an important issue when using technology. Project participants recommended the JISC TechDis publications such as the Senior Management Briefing 1 for details of how to increase accessibility. The Web site [www.techdis.ac.uk](http://www.techdis.ac.uk) provides more information.

## 1.10 Case studies of technology meeting stated goals

These are some of the case studies from the recommended literature. The cases demonstrate the many different goals that can be met by technology-enhanced assessment, over and above testing students against intended learning outcomes and delivering timely, tailored feedback. Please see page 4 for a description of the evidence categories.

**Table 3**

*Case studies of using technology to meet specific goals*

| Goal   | How others have done it   | Source and evidence category                          |
|--|---|---|
| Improve learning                               | Electronic voting systems and multiple choice questions   | Draper (2009) (3)                                     |
|  | Interactive engagement (a variety of technology-enhanced assessment and teaching methods)               | Hake (1997) (1a)                                      |
|  | Digital audio feedback, Leeds Metropolitan case study   | JISC (2009) (4)                                       |
|  | Online communication tools: discussion boards for peer and self-assessment, REAP project case study     | JISC (2007) (4)                                       |
| Learner self-regulation                        | Various technology used on two courses: French and Psychology   | Nicol (2009) (1a)                                     |
| Motivate students to study throughout the year | Phase assessment  | Marriott (2009) (2)                                   |
| Save time marking assessments                  | Automated marking   | JISC et al (2008) (4), Cassidy and Gridley (2005) (2) |
| Deliver feedback to large cohorts quickly      | Computerised feedback to open questions on a medical progress test, University of Dundee Medical School | In Mitchell et al (2003) (1a)                         |

| Goal  | How others have done it   | Source and evidence category                                 |
|---|---|--|
| Increase the quality of feedback  | Personalised learning resources provided according to assessment performance, Oldham College case study   | JISC (2007) (4)  |
| Quicker and better feedback   | Audio feedback, podcasting  | Rotheram (2009) (4)  |
| Provide the human touch   | Podcasts, audio feedback, Duckling Project, University of Leicester   | Seminar series (see Appendix D) (4)                          |
| Diagnose IT skills  | Online testing and feedback   | Sieber (2009) (2)  |
| Produce authentic questions   | Reproducing radiographs in full colour on screen, University of Nottingham Medical School   | JISC et al (2008) (4)  |
| Authentic assessment of skills needed in business and professional practice | Use of open-book, open-web exams for authentic assessment of learning outcomes  | Williams and Wong (2009) (2)                                 |
| Relevant assessment   | For example e-portfolios, reflective diaries, blogs or virtual world scenarios, Anglia Ruskin case study  | JISC (2007) (4)  |
| Any time, any where, on-demand assessments                                  | Online formative and summative assessments, Coleg Sir Gar case study  | In Whitelock (2006a); also JISC (2007) (both 4)              |
| Assessment at work for the equine industry                                  | Online assessment delivered via a commercial ISP, Warwickshire case study   | In Whitelock (2006a) (4)                                     |
| Keeping up to date with the latest professional guidelines                  | Formative e-assessment with feedback kept up-to-date using a standard word processing editor, Chesterfield Royal Hospital case study  | In Whitelock (2006a) (4)                                     |
| Evidence of professional practice   | E-portfolios, Chef's Course at West Suffolk College used digital photographs  | In Whitelock (2006a) (4), Newcastle Seminar (Appendix D) (4) |
| Enabling groups to work together  | Using WebPA for peer assessment of contributions to group work  | Gordon (2008), (3) Loddington (2009) (3)                     |
| Encourage deeper reflection   | Confidence or certainty-based marking i.e., students have to indicate how certain they are about each answer in an assessment using a numerical scale, medical school at University College London (UCL) case study, UCL and Glamorgan case studies | JISC (2007) (4)<br>In Whitelock (2006a) (4)                  |
|   | Assessed blog: students take turns to blog about a text. Remaining students comment on blogger's analysis through questions, pointing out omissions, correcting mistakes, challenging preconceived ideas. Exercise discussed in tutorials.          | University of Edinburgh case study in JISC (2009) (4)        |

| Goal  | How others have done it                                     | Source and evidence category |
|---|---|------------------------------|
| Aligning assessment, learning and teaching                        | JeLSIM simulation toolkit and PASS-IT assessment engine     | Ashton and Thomas (2006) (3) |
| Wide geographical dispersion, large numbers and on-demand testing | Online assessment, Calibrand British Citizenship case study | In Whitelock (2006a) (4)     |

## 1.11 What has limited the take-up of technology to meet goals?

As technology can be used to enable assessment and feedback to meet so many goals, it might be surprising that the approach has not been adopted by more HEIs.

### 1.11.1 First attempts

Warburton (2009) undertook research (evidence category 2) into why technology-enhanced assessment has not been implemented on the large scale that specialists in the field expected. He found that:

*“CAA [computer-assisted assessment] was widely acknowledged to offer the potential of productivity gains in terms of more efficient authoring, publication, delivery, marking and reporting, which was summed up by some respondents as an effective reduction in paperwork (Warburton and Conole, 2005).*

*However, it also emerged that where unsupported tutors sought these ‘quick wins’ without investing in preparative activities such as seeking the advice of experienced colleagues or setting up formative exercises and practice quizzes, the degree of risk taken on all at once could be so significant that colleagues were effectively discouraged from using CAA themselves.”*

Warburton also found a staged approach which he called “slow burn”. This approach was:

*“Where tutors aimed primarily for pedagogical improvements they took on less risk and their resultant trajectories were characterised by routine use of small scale quizzes with an initial emphasis on low stakes testing such as formative and diagnostic applications.*

*This sometimes progressed towards higher stakes testing on a larger scale. A staged approach was encouraged by learning technologists who recognised the value for tutors of learning to use complex CAA tools in less critical applications. High stakes applications such as examinations were seen by both learning technologists and experienced CAA using tutors as the final goal of CAA trajectories rather than a starting point.”*

Warburton described the ideal uptake “trajectory” for both institutions and tutors as a step-by-step approach through the following stages and goals:

- Planning;
- Piloting;

- Low stakes, small scale application;
- Low stakes, large scale application;
- High stakes;
- Embedded.

Section 10, Feedback to Stakeholders, provides information on the issues Warburton found different groups focused upon in relation to the “successful” uptake of technology-enhanced assessment.

### **1.11.2 Challenging a barrier to take-up: Does using technology disadvantage students?**

Other reasons for the limited take-up of technology-enhanced assessment lie in the perceived barriers that the technology presents. Recent research by John Dermo at the University of Bradford challenges some of the reasons put forward by academics for not using technology-enhanced methods, e.g., that technology can disadvantage older students or female students, that it is stressful for students to use, and is distrusted by many of them.

Dermo (2009) conducted a survey of 130 students at the University of Bradford. These students had experience of formative (low stakes) and summative (high stakes) e-assessment using multiple-choice and short answer question types. The formative tests provided feedback immediately or after a short time and the summative assessments provided no feedback, just the grade. Dermo cautions that “we cannot claim to have discovered universal truths, and must make our tentative generalisations based firmly on our own context.” However, his findings are based on collated statistics from Likert scales which he suggests makes the data “pseudo-qualitative or even quantitative”. We have categorised the data as 1b. His findings question some of the assumptions that stakeholders may have regarding technology-enhanced assessment.

Key findings were:

- *“Affective factors: students definitely did not think that e-assessment added to stress and there was an expectation that online assessment would happen at university.*
- *Validity: e-assessment was seen to be suitable for the students’ studies and able to deal with the complexity of their studies.*
- *Practical issues: health and safety were definitely not seen to be a problem and e-assessments were considered accessible.*
- *Reliability: e-assessment was seen to improve reliability of marking.*
- *Security: grades, logon and e-assessment in general were all considered secure.*
- *Teaching and learning: students believed that e-assessment was not just a gimmick but added value to learning, going hand in hand with e-learning, and feedback from e-assessment also helped learning.”*

This study also found that there was no significant correlation between age and student ratings. Dermo comments, “This may come as a surprise to those who assume that older students tend to view online activities less favourably than their younger fellow students.” The research also found that there was no significant difference in response between male

and female students. Dermo comments: "Again, this may come as a surprise to many, who might assume gender differences to figure more prominently".

The student concern that was raised during this research was that tests that use random-selection of items from a question bank are unfair. Dermo suggests that "using item analysis to check the difficulty level of the items in the bank" can address this concern.

Lee and Weerakoon (2001) also found that the technology did not disadvantage students. Their research (evidence category 1a) found that 61% of students had no anxiety about using a computer for assessment, despite their inexperience with tests delivered in this format. The students in the study were familiar with using computers for word processing but had a low level of experience for other uses (44% had moderate experience and 38% had high experience of word processing). The study found:

*"There was no correlation between student performance on the computerized test and any of the measures of computer anxiety (either total anxiety or interactive computer learning anxiety) or computer experience. This does not necessarily mean that computer experience or anxiety have no influence on test performance, but, rather, if they do, the influence is weak or masked by other factors, such as subject knowledge. The lack of correlation may also reflect the user-friendliness of the computerized LXR test software."*

The researchers also found that:

*"In this study there was also no significant difference between rankings, when students were ranked according to their total score in the paper test and according to their total score in the computer test."*

In this study, students sat a paper-based version of the test and then a computerized version. The researchers found that the cohort performed better on the paper test than the computer test. Lee and Weerakoon concluded that "computer-based multiple-choice questions assessments can be used with confidence for ranking, but caution is advised in using them for grading."

This finding is challenged, however, by the work of Ashton, Beevers, Korabinski and Youngson (2005) at Heriot-Watt. Their paper provides evidence from category 1a research into the effect of the medium in computer-aided assessment of school chemistry and college computing national examinations. Nearly 200 students sat either a traditional paper, a computerized version of the paper test, or a screen shot of the computerised version. They found that:

*"There is no medium effect when specific traditional paper examinations in Chemistry and Computing are transferred into electronic format. The effect of rewording for computer-delivered test questions [was] also investigated and again the conclusion is that no evidence of a difference could be found."*

Similar findings were made by Cassidy and Gridley (2005) who compared student test anxiety for paper-pencil and online summative tests. Their study (evidence category 1a) examined two successive year groups that had the same course experience except for the format of the summative assessment. They reported that, "The only meaningful difference reported by students in the two testing conditions was the heightened level of perceived threat reported by students taking tests on paper." The authors proposed that this outcome

is “mostly influenced” by the lack of personal control over the paper test. The online test was available over a week, with students being able to choose when they took it. The authors also noted that their students had experience of using online tools for research before taking the test.

### **1.11.3 Challenging a barrier to take-up: Can technology address higher order learning skills?**

As it is easier to find examples of technology being used for simple question formats, such as multiple-choice, and for questions that test the lower levels of Bloom’s taxonomy of educational objectives. Some academics query whether technology can be used to address higher learning skills effectively. This barrier to take up is challenged by authors working with more complex questions in a technology-enhanced environment. For example, Jordan and Mitchell (2009) and Butcher and Jordan (2010) discuss the merits of different tools for free-text technology-enhanced assessment. They conducted a study using a tool called FreeText Author (developed by Intelligent Assessment Technologies and the Open University) for short-answer free-text questions with tailored feedback. This study (evidence category 1a) found that the “answer matching has been demonstrated to be of similar or greater accuracy than specialist human markers.”

## **1.12 Summary of using technology to meet learning goals**

This section explored the different goals or objectives that technology enhancement can enable HEIs to meet through assessment and feedback, beyond testing student performance against intended outcomes.

### **1.12.1 Key messages**

There is evidence in the recommended literature that technology-enhanced approaches can enable HEIs to:

- Provide assessment and feedback that meets the principles and conditions for effective assessment, such as providing prompt, detailed feedback tailored to the student’s learning needs and offering frequent, on-demand formative tests. Without the technology enhancement, constraints such as tutor time, student numbers and geographical or temporal distribution, and costs can make it very difficult to meet all the criteria of effective assessment and feedback.
- Replace the tutor’s or another method’s role (such as paper-based marks and comments) in an integrated assessment and feedback process, so freeing the tutor’s time to devote to other activities.
- Add a dimension to the student experience that was not possible without technology, such as audio feed forward and feedback that students can work through in their own time and which adds the ‘human touch’ while potentially saving tutor time.
- Provide regular, online testing and rapid, detailed feedback which improves student performance, learning and self-regulation.
- Encourage students to stimulate and scaffold each other’s learning and increase social engagement, student collaboration, peer assessment, and retention through media such as discussion boards and group work.
- Create assessment that is authentic to the competence being tested, for example by providing audio and visual information that students would use to complete the task

in real life or requesting suitable audio and visual evidence of how a task has been performed.

- Support tutors and learning assistants to write effective assessments and feedback in a timely manner.
- Provide quicker and similar or more accurate marking than human markers.
- Offer students greater flexibility over the timing and location of their assessments.
- Ensure greater accessibility than for paper-based tests.

The recommended literature also provided evidence that challenges some of the barriers to the take-up of technology-enhanced methods. For example, studies which showed that, at least in the context of the study:

- Using technology does not disadvantage older or female students and it is not more stressful to use or distrusted by students, although the random-selection of test items was queried by students as being possibly unfair.
- There is no medium effect between paper and computer-based exams.
- Answer matching has been demonstrated to be of similar or greater accuracy than specialist human markers.

Reaping the potential benefits of technology enhancement does depend on the assessment and feedback being well designed. Design issues are the focus of the next section.

### **1.12.2 Underlying drivers**

The literature has many case studies of successful use of technology-enhanced assessment. Of the references that were recommended to us, 72% involved case studies or experienced practitioner views that appear in peer-reviewed journals or other reputable sources. A common feature of these success stories are the underlying drivers for the introduction of the technology-enhanced method. Usually technology-enhanced assessment and feedback needs to address at least one of the following drivers to be successful (Gilbert & Gale, 2007):

- Large classes;
- Stable content;
- Significant risk;
- Significant cost;
- Lower levels of Bloom's taxonomy;
- Geographical and/or temporal dispersion of students.

### **1.12.3 Practical issues**

The literature we were recommended includes articles on how important it is to consider practical issues (such as the provision of suitable PCs) when considering the feasibility of any technology-enhanced project. However, for the purposes of this report we have placed such practical issues under Section 6 Delivery.

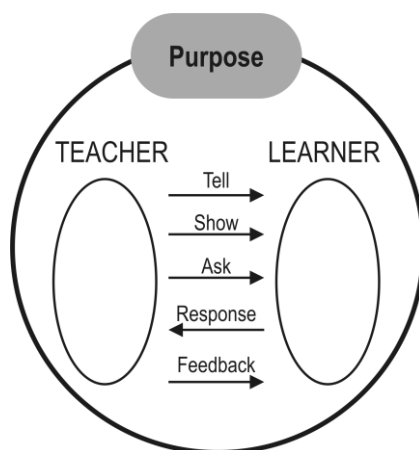
## 2. Designing assessment and feedback with technology enhancement

### 2.1 Introduction

In this section we focus on how to design effective technology-enhanced assessment and feedback so that it meets the needs of the:

- learning model adopted for course;
- learning outcomes;
- students;
- academics;
- other stakeholders.

Part of the design task is to take the course content outline and develop the learning model for the course. For example, the 'conversational' model of Laurillard (1993) may be abstracted to yield a model of an 'atomic' learning transaction, here called an instructional transaction, and illustrated in Figure 3 (Gilbert & Gale, 2007). As the figure shows, assessment and feedback is an integral part of the instructional transaction. Part of the assessment and feedback design task is therefore to develop the assessment elements in the 'conversations' or transactions which will achieve the desired learning outcomes.



**Figure 3:** *The Instructional Transaction*

Readers will notice that, with Section 1 Goals, this section forms the bulk of the report. This is because the literature recommended to us concentrates on what technology enhancement methods can deliver and how they can be used effectively. So, there are many case studies, literature reviews, and experienced practitioner guidance on effective design. For example:

*“Something I always try to point out is that it is intellectually a lot more stimulating to design an online or e-assessment than it is to sit there marking 300 scripts.”*

Professor Don Mackenzie, Senior e-Assessment Manager, Innovation 4 Learning, University of Derby, quoted in JISC (2007).



The 'Assessment for Learning' movement, which has had a significant influence on educational practice in the UK, emphasises the role of formative assessment in improving learning. Unpacking what is meant by formative assessment reveals that the learning gains which are obtained are also dependent on the feedback given to students. Findings from Winne et al (2006) (evidence category 1b) and the work of Boud (2002, evidence category 4) all stress the importance of finding 'devices' for self-monitoring and providing students with feedback about how they can study to learn, while Nicol & Macfarlane-Dick (2006, evidence category 4) have shown that formative assessment and feedback can help learners take control of their own learning.

To design effective feedback that achieves these goals and those listed in Section 1, it is important to write the assessment task and the feedback at the same time. We have therefore included feedback in this section. For references on writing effective questions, please see Section 3 Creation.

## **2.2 Design should be driven by pedagogy rather than technology**

Draper (2009) (evidence category 3) comments that "As with all educational technology, whether learning benefits are achieved depends not on the technology but on whether an improved teaching method is introduced with it." It is the pedagogy that drives the success of technology-enhanced methods rather than the technology itself. The educational benefits of using questions that were delivered using traditional means should still be achieved when they are delivered using technology. The use of technology may also deliver benefits of "speed, ease, and the capability to handle very large classes" (Draper, 2009).

We have included lists of design principles in Appendix C that apply to all forms of assessment and feedback, regardless of medium. This is because academics designing for technology-enhanced methods need to build on the effective practice in assessment and feedback gained using other methods. This effective practice can then be adapted for and make best use of the technology chosen.

The experts consulted for the *Report on Summative E-Assessment Quality (REAQ)* (Gilbert et al, 2009) identified pedagogy, i.e., mapping to intended learning outcomes, as the key characteristic of quality assessment after reliability and validity. This report is in evidence category 4. Report contributors also ranked 'more authentic' assessment, assessing a range of skills (such as offering hints and partial credits or asking for confidence factors), and accurate and helpful feedback, as key features of quality assessment.

The opportunities for more authentic assessment design are also highlighted by JISC's 2007 report on *Effective Practice with E-Assessment* (evidence category 4):

*"Evidence suggests e-assessments can provide assessment experiences that are more authentic – through the use of e-portfolios, reflective diaries, blogs or virtual world scenarios for example."*

## 2.2.1 Some examples of learning designs that use technology enhancement successfully

An example of how a successful learning design can be enhanced by technology is provided in Crouch and Mazur (2001), whose paper describes the results of ten years' experience of improved student results, (compared with traditional instruction and therefore in evidence category 1b) using a method they call Peer Instruction:

*"A class taught with PI [Peer Instruction] is divided into a series of short presentations, each focused on a central point and followed by a related conceptual question [MCQ example given]. Students are given one or two minutes to formulate individual answers and report their answers [using a poll] to the instructor. Students then discuss their answers with others sitting around them; the instructor urges students to try and convince each other of the correctness of their own answer by explaining the underlying reasoning. Finally, the instructor [...] polls students for their again (which may have changed based on the discussion), explains the answer and moves onto the next topic."*

They found that the "vast majority" of students who changed their vote after the peer discussion moved from an incorrect answer to the correct answer.

Draper (2009) (evidence category 3) discusses how this technique can be used with an electronic voting system, a technology used to display the question, capture the student responses, and display the votes for each option as a graph.

## 2.3 Designing assessment questions and feedback for technology enhancement

In this section we will look at the key messages from the recommended literature concerning question types that work well with technology-enhanced methods.

### 2.3.1 Designing multiple-choice questions

A wide-spread use of technology-enhanced assessment is the multiple-choice question (MCQs) test. These assessments are often where academics start using technology-enhanced methods. There are many software tools available to support tutors offering MCQs and variations on the multiple-choice theme such as true/false, yes/no, points on a scale, matching, ranking, fill-in-the blanks, and similar. Many in higher education associate MCQs with the "lowest kind of learning" (Draper, 2009), often because they have seen these questions used to test disconnected 'facts'. Sometimes, this may be what is required. "Objective tests [...] are well established in subject disciplines that depend on accurate recall of knowledge" (JISC, 2007, evidence category 4). Examples in the literature cover subjects such as medicine, dentistry, science, and law, where accurate recall of large quantities of knowledge is required.

Various authors have raised concerns that MCQs can only be used to test 'surface learning' or 'shallow learning'. This is in contrast to 'deep learning' that Draper (2009) describes as being "characterised by learning multiple relationships between items that can support multiple uses (i.e., transfer tasks) better." In this paper he suggests ways of giving MCQs

greater 'pedagogical power' (in addition to the peer interaction electronic voting system example given above) such as using:

- Assertion-reason questions that make an accurate statement then ask students to select the reason that explains the statement correctly.
- Asking learners to generate reasons for and against each option in an MCQ.
- Confidence-based marking.
- Asking students to create their own MCQs.

### **2.3.2 Open questions**

Technology enhancement is not just applied to MCQs, however. Jordan and Mitchell (2009) (category 1a) provide evidence for moving beyond the MCQ and using technology to ask open questions. They suggest that open questions are suitable for computerised delivery and feedback "if correct answers can be given in short phrases or simple sentences and the difference between correct and incorrect answers is clear-cut." Whitelock & Watt (2008, category 3) illustrate this using the Open University's 'Open Comment' system.

### **2.3.3 Moving beyond simple right or wrong answers**

The lower levels of Bloom's taxonomy are where many academics start to apply technology-enhanced assessment methods. However, there is evidence in the literature of how academics have moved into higher level learning outcomes and more subtle assessment questions when applying technology.

Ashton et al (2006) provide evidence (category 1a) that technology-enhanced methods can be used to mirror tutor marking practices in mathematical examinations. They explain how software was developed and some questions were redesigned to allow partial credits to be awarded and mathematical expressions to be entered by students in automated exams. This work was undertaken as part of the PASS-IT project.

Boyle and Hutchinson (2009) address the issue of whether or not sophisticated tasks can be assessed using technology. In this paper (category 3) they suggest that, "e-assessment will become an important and widely-used feature of education systems in the near future. Further, the types of questions and tasks used in near-future e-assessment may well be quite different from questions and tasks used in on-paper assessment, and in early implementations of computerised assessment."

## **2.4 Designing assessments for the contemporary learner**

Assessment tasks and questions are likely to develop in the near future to meet the needs of the contemporary learner. Elliott (2008) (category 4) suggests a way of modernising assessment that he calls 'Assessment 2.0'. This type of assessment responds to the learning styles of "digital natives" (a phrase coined by Prensky in *Digital Natives, Digital Immigrants*). Elliott lists the common characteristics of these learners from the literature:

- *skilled use of tools;*
- *active learning rather than passive receiving of knowledge;*
- *authentic learning experiences rather than contrived tasks;*

- *task (not process) oriented;*
- *just in time learning;*
- *search not memorise;*
- *doesn't know answer but knows where to find it;*
- *Google not libraries;*
- *collaborate not compete."*

He suggests that Assessment 2.0 "embraces the Internet and, more specifically, Web 2.0 [...] by using the same tools and techniques that students use at home and we use in the workplace." He identifies the following characteristics for Assessment 2.0 to match the needs of the contemporary learner:

- *"Authentic: involving real-world knowledge and skills.*
- *Personalised: tailored to the knowledge, skills and interests of each student.*
- *Negotiated: agreed between the learner and the teacher.*
- *Engaging: involving the personal interests of the student.*
- *Recognises existing skills: willing to accredit the student's existing work.*
- *Deep: assessing deep knowledge – not memorisation.*
- *Problem oriented: original tasks requiring genuine problem solving skills.*
- *Collaboratively produced: produced in partnership with fellow students.*
- *Peer and self assessed: involving self reflection and peer review.*
- *Tool supported: encouraging the use of ICT."*

Elliot suggests that the proof of competence "that best fits this form of assessment" would be:

- *"naturally occurring: already in existence or generated out of personal interest;*
- *multimedia: existing in text, audio and video format;*
- *digital: such as e-mail, instant message logs, blog posts, wiki contributions, audio and video recordings;*
- *distributed: may be scattered across various sources (such as web sites, blogs, inbox, iPod)."*

## **2.5 Designing the use of e-portfolios**

Using an e-portfolio learning process (regardless of the tool used) meets many of the needs of the contemporary learner provided above. Chang and Tseng (2009) conducted an experiment (evidence category 1a) into the effects of web-based portfolio assessment. This study was, "specifically tailored to the requirements of a computer application course in junior high schools." The researchers found that:

*"Use of the system has significant positive influence on students' performances. According to estimated effect size, the most significant indicators were reflection, self assessment, continuous improvement, goal setting, problem solving, data gathering, work and peer interaction."*

However, the authors found that peer-assessment performance was not enhanced significantly.

A different perspective on peer assessment was found by Boyle (2007) who presented the findings of a literature review (evidence category 3). Boyle's review includes the following points about e-portfolios:

- *“e-portfolio authoring encourages teachers and students to view drafts of work, and interact about them. The process of generating work is forefronted, rather than merely concentrating on the final product (Twining et al, 2006, p. 55).*
- *Tools in e-portfolios can allow teachers to ask students to upload materials at significant stages, thus illustrating what the students believe to be progress (an important element of self-assessment) (Twining et al, ibid.; McGuire et al, 2004, p. 4).*
- *Communications tools associated with e-portfolios can allow for the provision of varied feedback with respect to: authors (fellow students or teachers), formality, and mode of communication – writing or speech (McGuire, 2005, p. 267). Such variety can be useful for facilitating learning by students of different dispositions, experiences and cognitive styles.”*

To gain the benefits of using an e-portfolio process as evidenced above may require considerable work by tutors in designing and implementing the approach. This may be difficult when many HEIs have the goal of using technology to reduce workloads. However, when Strivens (2006) reviewed 13 cases of how portfolios are being used she found:

*“It was clear from the interviews that respondents were enthusiastic about portfolio assessment, often at the expense of their own extra time and effort in assessing. However it also became clear that well-established practice is still difficult to find and practitioners are engaging in trial and error in designing their systems. If portfolios are to continue to spread as an assessment method and fulfill their educational potential, more attention must be paid to the question of efficiency.”*

Strivens lists the efficiency measures that her respondents had developed to manage their workloads. In particular, the design of the e-portfolio assessment and feedback process involves:

- *“Spreading the assessment over time;*
- *Using self-assessment and peer-assessment (as well as tutor-assessment);*
- *Giving no feedback on the final assessment (but at key deadlines during course);*
- *Focusing feedback on specific items at final assessment;*
- *Most or all of the portfolio items are not graded or only graded as a Pass/Fail where the portfolio is a minor part of the module assessment;*
- *Specifying the organisation of documentation so the assessor can find it easily;*
- *Candidate presents/assessor selects a sample of evidence to assess;*
- *A robust and tested moderation strategy is used (an area that the author identifies as needing further work).”*

So, whether or not the potential benefits of e-portfolios are achieved and the challenges managed well, depends on the design used.

## 2.6 Designing technology enhancement for group work and peer assessment

A case for including group work in higher education courses is presented by Gordon (2008). His paper (category 3) examines the issues involved and suggests ways that technology-enhanced methods can be used to overcome some of the challenges. The author discusses his experience of group work on a computing course, “cooperative learning, where learners are required to work in planned groups working on structured tasks.” He identifies that:

*“One of the major concerns with students regarding group work [assessment] is that some students will not ‘pull their weight’. To attempt to ensure that a group assignment is successful, it is important to set clear learning outcomes which identify that group coordination and work is part of the marking scheme and to indicate to students the weighting attributed to this. There are numerous approaches to how a group mark is assigned to students – one ideal approach is simply to assign the group mark to all the group members. This can be perceived as fair, but does not allow for any differentiation in input by the individual group members – something that does concern many students.”*

To assess the different contributions of individual students, Gordon recommends the use of peer assessment. “Peer assessment offers a number of benefits to both the teacher and the students. One particular benefit for students in using peer assessment is that students are required to consider and apply marking criteria. This is particularly useful for first year students who may be unfamiliar with the kinds of marking criteria used in the U.K. – especially students who have come through non-traditional routes or from overseas. For the teacher, it can reduce the marking burden”. He suggests three different ways of using peer assessment:

- Each group must agree on a weighting of contribution to the group submission (tutors will have to mediate any disagreements);
- Each student assesses each student in the group and possibly themselves (large amount of data to collate);
- Groups mark other groups’ outputs such as a presentation (again a large amount of data needs to be collated).

Gordon discusses how WebPA, a tool developed by Loughborough University, can automate many of the aspects of peer assessment of group work. Gordon reports:

*“This tool allows teachers to set up modules and groups of students within these modules, and provides an interface for students to mark their peers within their group against the criteria defined by the teacher. The impact of these marks provides a weighting that is used to attribute the weighted version of the overall group mark to the individual students. The teacher has flexibility to determine what proportion of the group mark is subject to the peer weighting.*

*This system provides an efficient way to manage the peer assessment process, and provides the anonymity in input to give students the freedom to potentially give truer assessments of their peers. Whilst requiring a separate installation and*

*systems support issues the practical value for large classes makes it a worthwhile technology to investigate.”*

## 2.7 Designing feedback for a technology-enhanced environment

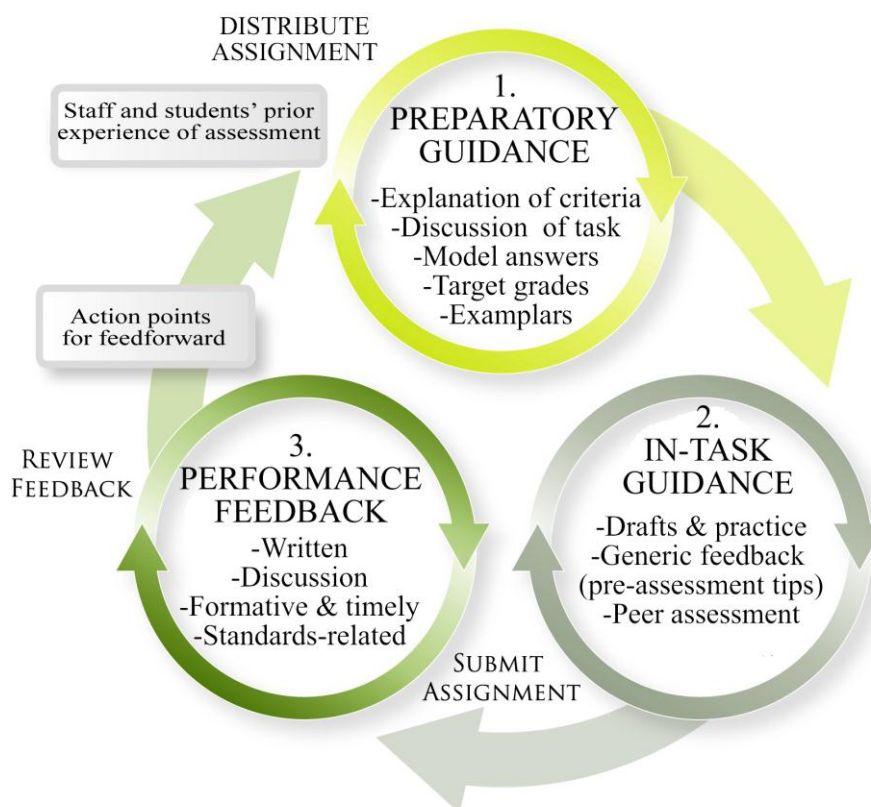
This subsection looks at the evidence for specific characteristics of feedback and feed-forward.

### 2.7.1 The case for developing appropriate feedback

When Maclellan (2001) conducted a survey (evidence category 2) of staff and students at the University of Strathclyde on their perceptions of assessment for learning, students' perceptions were that the feedback they received was not “routinely helpful in itself or a catalyst for discussion.”

This finding is supported by the results of the National Union of Students (NUS) surveys. Gilbert et al (2009) quote the NUS's principles of effective feedback which include that it “should consist of effective and constructive feedback”. The NUS comments that “effective feedback on assessment is a crucial aspect of assessment processes and a key feature of enhancing the learning process.”

Beaumont, O'Doherty and Shannon (2008) present the following diagram to illustrate how feedback and feed forward operate as a dialogic cycle:



**Figure 4**

*Dialogic feedback cycle, Beaumont et al (2008)*

Further evidence for the role of feedback in an assessment strategy is provided by Whitelock (2010, evidence category 4) who points out that, “Formative assessment assists the on-going learning cycle while summative assessment is not cyclical and assists with ascertaining the progress made by students at a particular moment in time e.g., at the end of a course.” Feedback from frequent formative assessment is therefore a vital component of a course’s learning strategy.

Specific recommendations for the design of feedback are given by Gibbs and Simpson (2004, evidence category 3). These recommendations are presented as a number of conditions that must be met if assessment is to support learning, see Appendix C for the full list. Seven of these conditions relate specifically to feedback. Assessment will support learning if feedback:

1. Is sufficient (in frequency, detail).
2. Is provided quickly enough to be useful.
3. Focuses on learning rather than marks.
4. Is linked to assessment criteria/expected learning outcomes.
5. Makes sense to students.
6. Is received by students and attended to.
7. Is acted upon to improve work and learning.

In addition to these conditions of effective feedback, Sadler (1989) specifies that effective feedback that promotes learning enables students to identify:

1. What good performance is (the goal or standard set).
2. How current performance relates to good performance.
3. How to close the gap between current and good performance.

Nyquist (2003), offers even more detailed recommendations. He reviewed 185 research studies on the effects of feedback in higher education (evidence category 4 as the study is unpublished), and developed a typology of different kinds of formative assessment:

**Table 4**  
*Typology of formative assessment*

|                               |  |
|-------------------------------|--|
| Weaker feedback only          | Students are given only the knowledge of their own score or grade, often described as ‘knowledge of results’.  |
| Feedback only                 | Students are given their own score or grade, together with either clear goals to work towards or feedback on the correct answers to the questions they attempt, often described as “knowledge of correct results”. |
| Weak formative assessment     | Students are given information about the correct results, together with some explanation.  |
| Moderate formative assessment | Students are given information about the correct results, some explanation, and some specific suggestions for improvement.   |
| Strong formative assessment   | Students are given information about the correct results, some explanation, and specific activities to undertake in order to improve.  |



Nyquist calculated the average standardized effect size for the studies for each type of feedback. These numbers show that “unless feedback engages the students in mindful activity, little or no learning will result, no matter how much time and care is spent by assessors in crafting that feedback”, (Wiliam, 2008).

Further evidence for well-designed feedback having a positive effect on achievement is provided by Hattie’s (1987) meta-analysis of student learning (evidence category 1b):

- “Overall, the effects of feedback were correlated with subsequent achievement and confirmation and review feedback had a larger effect than correct-response feedback” (Schmitt, 1983);
- “Diagnostic feedback had a higher relationship with achievement [than remediation]” (Yeany and Miller, 1983).

### **2.7.2 Designing feedback: good practice**

The Academy’s project *Student Enhanced Learning through Effective Feedback (SENLEF)* completed in June 2004 produced a report that includes case studies focusing on the principles of effective feedback practice, (Juwah et al, 2004, evidence category 4). Two of the authors of this report, David J. Nicol and Debra MacFarlane-Dick, wrote an article in 2006 on *Formative assessment and self-regulated learning: a model and seven principles of good feedback practice*. This article focuses on how tutors can “help students take control of their own learning, i.e., become self-regulated learners” by providing appropriate feedback. The authors explain their key argument as “students are already assessing their own work and generating their own feedback, and that higher education should build on this ability.” Nicol and MacFarlane-Dick put forward seven principles of feedback to promote self-regulated learning, each supported by evidence from research.

To define self-regulation the authors quote Pintrich and Zusho (2002):

*“Self-regulated learning is an active constructive process whereby learners set goals for their learning and monitor, regulate and control their cognition, motivation and behaviour, guided and constrained by their goals and the contextual features of the environment.”*

One of the main external features is Nicol and MacFarlane-Dick feedback provided by the tutor or peers in person or via technology. They propose:

*“Good quality external feedback is information that helps students troubleshoot their own performance and self-correct: that is, it helps students take action to reduce the discrepancy between their intention and the resulting effects.”*

#### **Seven principles of good feedback practice**

Nicol and MacFarlane-Dick characterise good feedback practice as:

1. *“Helps clarify what good performance is (goals, criteria, expected standards);*
2. *Facilitates the development of self-assessment (reflection) in learning;*
3. *Delivers high quality information to students about their learning;*
4. *Encourages teacher and peer dialogue around learning;*
5. *Encourages positive motivational beliefs and self-esteem;*

6. *Provides opportunities to close the gap between current and desired performance;*
7. *Provides information to teachers that can be used to help shape teaching.”*

Their article discusses the evidence for each principle and how the principle can be applied in practice. For example, for principle 1 the authors suggest providing students “with written documents containing statements that describe assessment criteria and/or the standard that define different levels of achievement” complemented by materials that clarify what these statements mean such as exemplars of the performance needed, peer marking and review.

Whitelock (2010, category 4) suggests that feedback should also include “advice for action”, i.e. what the student should do next to bridge the gap between current and desired performance and/or advice to use in future learning tasks and assessments.

Further evidence that information to move students from their current to the desired performance is the key for the effective use of e-feedback is provided by Dexter (2010). In this category 3 paper Dexter presents some findings from the design and evaluation of software known as ETIPS (educational theory into practice software). It was developed to provide K-12 educators with case studies where they could practise using concepts from their university education courses in a classroom and school setting. A ‘student model’ is used which defines the declarative, procedural and contextual knowledge that students are being tested on with e-feedback explaining how the student can improve their performance. It is vital that students perceive that the knowledge and skill captured in the student model is important, so that they respond to the e-feedback with further learning.

### **2.7.3 Designing audio feedback**

As well as the general principles of feedback provided above, the references recommended to us included specific advice for designing audio feedback. Middleton and Nortcliffe’s (2010) study was designed to identify factors which affect the implementation of audio feedback. Semi-structured interviews were conducted with academics about their use of audio feedback (evidence category 3) and the authors have produced a set of audio feedback design principles from their findings. These principles are that audio feedback should be:

1. *“Timely and meaningful.*
2. *Manageable for tutors to produce and the learner to use.*
3. *Clear in purpose, adequately introduced and pedagogically embedded.*
4. *Technically reliable and not adversely determined by technical constraints or difficulties.*
5. *Targeted at specific students, groups or cohorts, addressing their needs with relevant points in a structured way.*
6. *Produced within the context of local assessment strategies and in combination, if appropriate, with other feedback methods using each medium to good effect.*
7. *Brief, engaging and clearly presented, with emphasis on key points that demand a specified response from the learner.*
8. *Of adequate technical quality to avoid technical interference in the listener’s experience.*

9. Encouraging, promoting self esteem.
10. Formative, challenging and motivational.”

A general structure for the design of audio feedback is offered by Rotheram (2009, evidence category 4) which “may be useful in a variety of circumstances”:

- “Introduce yourself to the student in a friendly manner.
- Say which assignment you are giving feedback on.
- Outline the main elements of the comments which you’ll be giving (see below).
- Work steadily through the assignment, amplifying and explaining notes you’ve put in the margins and, especially at the end, making more general points.
- Refer to the assessment criteria.
- Explain your thought processes as you move towards allocating a mark.
- Give the mark (perhaps – see above remarks on student engagement and moderation).
- Offer a few (reasonably attainable) suggestions for improvement, even if the work is excellent.
- Invite comments back from the student, including on the method of giving feedback.
- Round things off in a friendly way.”

## 2.7.4 Case studies highlighting the design of technology-enhanced feedback

All these case studies are evidence category 4 unless other wise stated.

**Table 5**

*Case studies of designing technology-enhanced feedback*

| Project  | Technology used  | Details   |
|--|--|---|
| Technology, Feedback, Action!<br>Sheffield Hallam University   | <i>Blackboard:</i> to deliver feedback and grades<br><i>Assignment Handler:</i> students must submit a reflective account or action plan based on assessment feedback before grades are released.<br><i>Feedback Wizard:</i> linked feedback to assessment criteria. | Seminar briefing paper on the Academy’s website. Guides available from the Higher Education Academy EvidenceNet: <ul style="list-style-type: none"> <li>• Feedback: a student guide to using feedback</li> <li>• Using technology to help students engage with their feedback: A best practice guide for academic staff</li> <li>• Using technology to help students engage with their feedback: A ten minute guide for senior managers</li> </ul> See Appendix D for links |
| Podcasting in Assessment: New Technology in Higher Education Research (PANTHER)<br>University of Leicester | Podcasting   | Seminar briefing paper on the Academy’s website<br>Parent project website:<br><a href="http://www.le.ac.uk/duckling">http://www.le.ac.uk/duckling</a><br>See Appendix D for links   |

| Project  | Technology used  | Details   |
|--|--|---|
| Moving Forward Through Feedback<br>University of Reading                                   | Video feedback   | Seminar briefing paper:<br><a href="http://www.reading.ac.uk/asset/">http://www.reading.ac.uk/asset/</a>  |
| Embedding feedback and feed forward through curriculum design<br>University of Sheffield   | Use of VLE to deliver feedback and feed forward<br>Team approach to curriculum design                      | Seminar briefing paper on the Academy's website:<br><a href="http://search3.openobjects.com/kb5/hea/evidencenet/resource.page?record=SuRb2PddpCc">http://search3.openobjects.com/kb5/hea/evidencenet/resource.page?record=SuRb2PddpCc</a> |
| Improving Assessment and Feedback Practices in a Technology-Enhanced Teaching and Learning | Electronic Voting Systems, Tablet PCs, Interactive tablets, i-pods, mobile phones, in-class text messaging | Rapid Reaction and Response (R3)<br><a href="http://staffnet.kingston.ac.uk/~ku36708/RRR/">http://staffnet.kingston.ac.uk/~ku36708/RRR/</a>   |
| WebPA  | Peer assessment of contributions to group work   | <a href="http://www.webpa.ac.uk/">http://www.webpa.ac.uk/</a>   |
| Provide immediate feedback and hints to answer exam questions                              | Online exam with marks deducted depending on the level of help needed to get the correct answer            | Open University case study in Whitelock (2010)  |
| Active learning  | Voting systems using handsets or PDAs  | Whitelock (2010) example of MCQs on PDA, group discussion and decision on answer. Key is design of MCQ to reinforce or question what has already been learned.  |
| PASS-IT<br>Using online assessment to enhance teaching and learning                        | Online assessment of maths   | Ashton and Wood (2006) (Evidence category 3)  |

### 2.7.5 Designing feedback: when to give it

The best time to give feedback is when there is still time for the student to act upon it and improve their performance (Nicol and MacFarlane-Dick, 2006 and Sadler 1989). With some technology-enhanced methods immediate feedback is intrinsic to the skill being assessed, for example through computer simulation of the task.

Nicol and MacFarlane-Dick (2006) suggest the following strategies:

- *“Provide feedback on the work in progress and increase opportunities for resubmission.*
- *Give feed forward and group feedback.*

- *Set two-stage assignments: provide action points then involve students in groups identifying their own action points.*

Although tutors might offer feedback at a time when students can act upon it to improve their performance, students may not take the opportunity to do so unless encouraged. Whitelock (2010) cites previous authors (Wojtas, 1998; Weaver, 2006; Price and O'Donovan, 2008) who found that, although students ask for feedback they often seem only interested in their grade. Lindsay Jordan reports on efforts to engage students with all aspects of their feedback:

*“Asking students to respond to their feedback has been found to be a highly effective method of increasing distance learners’ engagement with summative feedback. Feedback response tasks encourage students to focus on how they will use their feedback in the future, while simultaneously allowing them to enter into a dialogue with the tutor on any points that require clarification.”* Jordan (2009, evidence category 3.)

### **2.7.6 Using peer assessment to give feedback**

A feedback approach that is likely to engage students is to include peer assessment. Nicol (2010, evidence category 3) discusses how peer assessment can be used to provide “feedback dialogue” and puts forward the following benefits:

*“Peer feedback scenarios where students receive comments on an assignment from many other students provide a richness and volume of dialogue that is difficult for a single teacher to match. In such situations, students must actively process and reprocess feedback input from a variety of sources and are potentially exposed to multiple levels of analysis and scaffolding.*

*[The] construction of feedback is likely to heighten significantly the level of student engagement, analysis and reflection with feedback processes.*

*[Further,] where peers generate and receive feedback in relation to the same assignment task (i.e. an essay that all students are writing), they learn not only about their own work but also about how it compares with productions of other students.”*

Nicol recognises that some students have a “lack of confidence in their peers and prior predispositions to solo working” and suggests that teachers comment on the peer comments when the peer working is first introduced to overcome these obstacles.

Evidence for the efficacy of peer assessment of group work has been discussed earlier in Section 2 Design.

## **2.8 Summary for designing technology-enhanced assessment questions and feedback**

In this section we explored the evidence in the recommended literature that shows:

- The pedagogy and therefore the design used is the most important component in the success of assessment and feedback with technology enhancement rather than the technology itself; technology is the enabler.

- Academics designing for technology-enhanced methods need to build on the effective practice in assessment and feedback gained using other methods. This effective practice can then be adapted for and make best use of the technology chosen.
- Technology can be used to design assessment and feedback that:
  - Is more authentic to the skills being assessed;
  - Can assess a range of skills and provide accurate and helpful feedback in a short period of time using designs such as hints, partial credits and asking the student how confident they are of their answer;
  - Is more accurate than alternative methods, for example using electronic voting instead of a show of hands;
  - Adds 'pedagogical power' to multiple-choice question, for example by using assertion-reason questions;
  - Asks open questions;
  - Meets the learning needs of the contemporary learner who prefers: active, personalised and just in time learning; authentic tasks; knowing where to search for information rather than memorising it; skilled use of tools and collaboration amongst other characteristics. This 'Assessment 2.0' makes use of the Internet, especially Web 2.0 to match assessment and feedback to the learning characteristics of students;
  - Encourages peer and self-assessment, student learning, and reflection through a process of e-portfolios. This design can also be used to reduce the workload placed on the tutor for marking and giving feedback.
- The design of marking schemes is very important to the success of group work. A successful approach is to use peer assessment supported by technology such as WebPA.
- Appropriate and constructive feedback plays a vital role in the success of assessment, especially assessment for learning. The literature offers detailed guidance on the design of effective feedback such as conditions, research-backed principles and a typology, as well as specific advice for the design of audio feedback.
- Good feedback helps students troubleshoot their own performance and self-correct. For example, one of Nicol and MacFarlane-Dick's principles for good feedback practice is helping to clarify what good performance is by providing appropriate documents e.g., the intended learning outcomes and competences for the course.
- Good feedback should also include 'advice for action' which tells students how to bridge the gap in their learning.
- Feedback has the most impact when there is still time for the student to act on it and improve their performance. Students may need encouragement to engage with the feedback comments rather than just the mark. One design approach that is likely to engage students is the use of peer assessment.

### 3. Creating the assessment and feedback using technology

#### 3.1 Introduction

The literature that was recommended to us rarely focuses on the mechanics of the technology involved, such as how to choose a computer-assessment tool or decide on a format for audio/video files. It focuses instead on how the design of the assessment promoted the competences being tested or how technology-enhanced methods overcame a problem in realising a learning design. Whitelock and Brasher (2006, evidence category 3) call for a “pedagogically driven model for e-assessment” rather than a “technologically and standards led framework”.

The message is clear that the design of the assessment comes first. However, the technology available is an important factor in what it is possible to create. In Section 2 Design of this report we have considered the evidence for using different types of technology from the simple (such as an electronic voting system) to the complex (such as business simulations) and the assessment designs that make them effective. Below we explore the characteristics of useful technologies.

Given the technology available, another important consideration is the people who will be involved in creating and delivering the assessment and feedback. It may be down to an individual tutor or to a team that might include some or all of the following roles: tutors, learning technologists, technology support staff and senior managers.

#### 3.2 Characteristics of useful technologies

We have seen evidence for effective assessment and feedback with simple tools that provide a limited range of question and response types, for example, quizzes in a VLE such as Moodle. Some HEIs have developed their own software tools more suited to their needs. Examples of these tools include the following.

**Table 6**  
*Examples of software tools developed by HEIs*

| Tool                      | HEI   | Reference                  |
|---------------------------|---|----------------------------|
| OpenMark<br>Open Comment  | Open University   | Whitelock (2006b)          |
| FreeText Author           | Open University and Intelligent Assessment Technologies | Jordan and Mitchell (2009) |
| TRIADS                    | Derby   | Whitelock (2006a)          |
| PASS-IT assessment engine | Heriot-Watt and the Scottish Qualifications Authority   | Ashton and Wood (2006)     |
| E-rater and C-rater       | USA   | Gipps (2005)               |

Building on the experience of creating an assessment system at Newcastle University, MacKenzie (2003) suggests that an ideal e-assessment system can:

- “Deliver pre-test and pre-question learning materials if required;
- Provide the widest possible range of question styles so that assessment design is not compromised by the limitations of the system;
- Deliver informative, detailed and context/score sensitive feedback to the student;
- Provide informative feedback on student and question performance to the tutor so that tutorial help may be targeted and improvements may be made for future runs;
- Provide a wide variety of question selection and sequencing options to facilitate deployment of assessments in a wide variety of delivery environments and modes of operation;
- Deliver summative assessment in a reliable and secure manner.”

Ashton, Beevers, Schofield and Youngson (2004, evidence category 3) build on the importance of technology-enhanced assessments providing suitable information for the tutor in their paper on informative reports. The reports they have created include ones that have visual information that provide useful information quickly, for example:

| Attempt | Duration | Q1 | Q2 | Q3   | Q4 | Q5 | Q6 | Q7  | Q8  | Q9  | Q10 | Q11 | Q12 | Q13 | Total Mark | Student Responses |
|---------|----------|----|----|------|----|----|----|-----|-----|-----|-----|-----|-----|-----|------------|-------------------|
| 1       | 00:33:46 | 2  | 4  | 3    | 0  | 3  | 2  | 4   | 3.5 | 2   | 0   | 1   | 3   | 0   | 27.5       |                   |
| 2       | 00:33:46 | 1  | 4  | 2.09 | 5  | 3  | 2  | 4   | 3.5 | 1.5 | 1   | 0   |     |     | 27.09      |                   |
| 3       | 00:39:48 | 2  | 4  | 0    | 0  | 3  | 2  | 2.5 | 3.5 | 2   | 3   | 4   | 4   |     | 30         |                   |
| 4       | 00:23:02 | 2  | 4  | 2.09 | 0  | 3  | 2  | 2.5 | 0   | 0   | 0   | 0   | 3   |     | 18.59      |                   |
| 5       | 00:01:28 | 2  |    |      |    |    |    |     |     |     |     |     |     |     | 2          |                   |
| 6       | 00:02:53 | 0  | 4  |      |    |    |    |     |     |     |     |     |     |     | 4          |                   |
| 7       | 00:01:46 | 0  |    | 3    |    |    |    |     |     |     |     |     |     |     | 3          |                   |
| 8       | 00:02:32 | 0  |    |      | 5  |    |    |     |     |     |     |     |     |     | 5          |                   |
| 9       | 00:01:47 | 0  |    |      |    | 3  |    |     |     |     |     |     |     |     | 3          |                   |
| 10      | 00:01:53 | 0  |    |      |    |    | 2  |     |     |     |     |     |     |     | 2          |                   |
| 11      | 00:01:51 | 0  |    |      |    |    | 0  |     |     |     |     |     |     |     | 0          |                   |
| 12      | 00:01:33 | 0  |    |      |    |    | 1  |     |     |     |     |     |     |     | 1          |                   |
| 13      | 00:00:56 | 0  |    |      |    |    | 2  |     |     |     |     |     |     |     | 2          |                   |
| 14      | 00:01:26 | 0  |    |      |    |    | 0  |     |     |     |     |     |     |     | 0          |                   |
| 15      | 00:00:39 | 0  |    |      |    |    | 2  |     |     |     |     |     |     |     | 2          |                   |
| 16      | 00:00:25 | 0  |    |      |    |    | 0  |     |     |     |     |     |     |     | 0          |                   |
| 17      | 00:01:39 | 0  |    |      |    |    | 2  |     |     |     |     |     |     |     | 2          |                   |
| 18      | 00:04:42 | 0  |    |      |    |    | 0  | 2.5 |     |     |     |     |     |     | 2.5        |                   |
| 19      | 00:02:33 | 0  |    |      |    |    |    |     | 3.5 |     |     |     |     |     | 3.5        |                   |
| 20      | 00:01:16 | 0  |    |      |    |    |    |     | 1.5 |     |     |     |     |     | 1.5        |                   |
| 21      | 00:01:15 | 0  |    |      |    |    |    |     | 2   |     |     |     |     |     | 2          |                   |
| 22      | 00:01:21 | 0  |    |      |    |    |    |     | 0   |     |     |     |     |     | 0          |                   |
| 23      | 00:02:17 | 0  |    |      |    |    |    |     | 1.5 |     |     |     |     |     | 1.5        |                   |
| 24      | 00:03:04 | 0  |    |      |    |    |    |     | 5   |     |     |     |     |     | 5          |                   |
| 25      | 00:00:49 | 0  |    |      |    |    |    |     | 2   |     |     |     |     |     | 2          |                   |
| 26      | 00:02:29 | 0  |    |      |    |    |    |     | 0   | 1   |     |     |     |     | 1          |                   |
| 27      | 00:01:29 | 0  |    |      |    |    |    |     |     | 1   |     |     |     |     | 1          |                   |
| 28      | 00:02:28 | 0  |    |      |    |    |    |     |     | 0   |     |     |     |     | 0          |                   |
| 29      | 00:01:36 | 0  |    |      |    |    |    |     |     | 0   |     |     |     |     | 0          |                   |

**Figure 5**  
Example report from the PASS-IT project

This report “shows that this student began by attempting the majority of the questions in the assessment, and then to repeatedly target specific questions until they were ready to move on (as each question was randomised the student could get different parameters to practise



with each time they sat a test). If teachers were only to review summary marks for an assessment a student may appear to be scoring very poorly whereas visual representations such as that shown [...] allow patterns like this to be identified.”

When it comes to complex technology-enhanced assessments such as business simulations, Thomas et al (2004, evidence category 3) suggest using integrated simulations to increase “quality and efficiency of assessment. [...] the ability to link them with an assessment engine will allow students to be assessed in the same environment in which they learn and will integrate the assessment activity more fully into the educational process.” The authors discuss the functionality and potential usage of prototype simulation software, in particular for the assessment of higher order skills.

Another aspect to consider when selecting an appropriate software tool is the issue of accessibility for students and tutors. Sloan and Walker (2008, evidence category 3) review tools for accessible authoring and conclude that “accessible web content [...] cannot be created by automated means alone, and authors must either develop awareness of accessible design, or be supported by the authoring tool.” Ideally, authors would develop this awareness while using an authoring tool that facilitates the creation of accessible assessments and feedback and highlights potential issues for concern. Authors wishing to develop this awareness are supported by the JISC TechDis publications and website.

An added dimension that technology can bring to the contemporary learner is highlighted by Whitelock (2010) who explains that, “Web 2.0 tools facilitate collaboration and offer the potential to move away from the assessment of individual skills to implement a social constructivist view of learning”.

### **3.3 Using a team approach to the creation of assessment and feedback**

A mixture of different skills is required to produce effective technology-enhanced assessments and feedback. Even the simplest technology requires tutors to develop knowledge of how to use it properly. The more complex the technology and the higher quality the outputs needed, the more support is needed from technical experts and learning technologists (Gilbert & Gale, 2007 evidence category 4).

This may raise concerns about the cost-effectiveness of technology-enhanced methods. As Gipps (2005, evidence category 3) comments, “those who are driving online teaching, learning and assessment in higher education cannot ignore the resource issues. Unless there are cost-effective ways of introducing high-quality ICT-based assessment, it will continue to be a minority activity.”

Resource issues are also highlighted by Whitelock (2006b, evidence category 4) who reports on the importance of the “creativity of staff, both academic and technical, to create formative e-assessments with systems which provide a restricted repertoire of questions and feedback to students.” She also identifies that “the development process is time-consuming and costly”.

The demands of using technology and therefore a new approach to the design of assessment and feedback mean that a team approach often results in the highest quality output. Gilbert et al (2009) investigated how to create good quality e-assessment by

interviewing experienced and successful practitioners, (evidence category 4). They recommended, “Use a team approach to the development of e-assessments to ensure that subject-matter expertise, e-assessment expertise and IT delivery expertise is included.”

### **3.4 Staff development and support needed when using technology**

As we saw in Section 1, Warburton (2009) points out that one of the reasons why technology has not been embraced for assessment and feedback in the volumes predicted is because inexperienced, unsupported tutors have applied technology enhancement for “quick wins” and had poor outcomes. Staff development and support is therefore a key factor in the success or failure of technology-enhanced assessment projects.

Zakrzewski and Steven (2000) also emphasise the importance of staff development for all academic staff involved in creating and using technology-enhanced assessment. This paper is based on their experience of implementing such systems (evidence category 3). In particular they highlight the need to “explore pedagogic issues surrounding the construction and use of objective tests” and to write a “test specification”.

Other aspects of staff development are explored by Conole and Warburton (2005). Their paper (evidence category 3) highlights the time staff need to learn the software and to develop high quality questions, as well as the need to develop staff skills and offer them support.

As well as developing effective questions, staff need support in providing effective feedback. Kochakornjarupong and Brna’s (2010, category 1a) empirical testing of McFeSPA systems revealed that the underlying rules in the feedback model would be helpful to the training of tutors in other domains since it would guide them in the production of effective feedback. These rules include providing:

- *“Knowledge of a range of feedback patterns, and what makes them appropriate for different situations;*
- *Knowledge of the kinds of feedback that can be deployed and when these should be used;*
- *Familiarity with being scaffolded and taking more responsibility for providing feedback as the system is faded;*
- *Knowledge of the kinds of errors and weaknesses that manifest themselves in the student’s answers.”*

Recognising the need to support tutors giving feedback, The Open University has developed a tool called Open Mentor (Whitelock 2010, evidence category 4). “This tool analyses and displays the different types of comments provided by the tutor as feedback to the student. It then provides reflective comments to the tutor about their use of feedback.” The system “guides tutors into providing clearer positive reinforcement for high achievers and prompts tutors to advise all students about how to improve their grades.” Whitelock et al (2003, category 3) provides more details.

Sim et al (2004, evidence category 3) advises that “focused staff development may help alleviate a number of issues, such as guessing, testing various cognitive skills, using appropriate question styles and accessibility.”

An example of the type of staff development that is appropriate is given by Walker, Adamson and Parsons (2004, evidence category 3) who detail online staff education for using online assessment:

*“The core material and assignments are delivered entirely online via Blackboard, the course runs for a five week period each semester and is led by elearning professionals from the University’s Learning Enhancement Unit with input from academic staff with significant experience in the field, adding a ‘practitioner’s perspective’. [...] Typically three assignments are to be completed each week which let participants put their new skills into practice.”*

Gilbert et al (2009, evidence category 4) make the point that it is not just internal staff who need support when using new technology. They recommend tutors “help external examiners understand how the e-assessment works and what the different questions that will be generated will contain.”

### **3.4.1 Sources of advice for staff designing individual questions and overall assessments**

Project participants recommended the following guides for authors writing technology-enhanced assessments. These are written by experienced practitioners and fall into evidence category 4.

- Ashton and Youngson (2004): Creating Questions for Automatic Assessment in Mathematics.
- CAA Centre Resources: Guide to objective tests.
- Elliott (2007) Guide to Writing Objective Tests.
- PASS-IT (2004): Good Practice Guide in Question and Test Design.
- Wills et al (2010): Writing effective e-assessments: Hints and tips from practitioners.
- Jordan (2009): Examples of online feedback for collaborative learning.

## **3.5 Summary for creating assessment and feedback with technology enhancement**

This section provided evidence from the recommended literature for:

- The characteristics of useful technologies to use for assessment and feedback.
- Taking a team approach to the creation of technology-enhanced assessment and feedback as successful implementation requires skills in the application of technology and how to use to the technology itself as well as learning and the subject content.
- The vital role of staff development and support when introducing and developing assessment and feedback with technology enhancement.

## 4. Testing

### 4.1 Key messages about the testing stage

From their experience of implementing technology-enhanced assessment, Zakrzewski and Steven (2003, category 3) recommend:

*“A pre-test of the examination will be required. The aim of the pre-test is to ascertain the reliability and validity of the examination. Pre-tests are carried out by the academic member of staff responsible for designing the examination and a member of the support services. The pre-test will clarify the suitability of the examination at the level intended [year of study], the technical integrity of the examination, the relevance of blind guessing and that 100% can be obtained. A pre-test event log will record the results of the pre-test and any actions that need to be taken. All CBA [computer-based assessment] examinations should be pre-tested before they are sent externally for moderation.*

*A formal system test is required prior to the examination event. The aim of the formal system test is to ascertain the reliability of the system. The examinations undergo a ‘dry run’ in the arena that will be used by the students. For Web-based systems, the system requires testing on all platforms and browsers to ensure comparability of performance and comparability of presentation. Every question is answered and the results checked to ascertain that they have been accurately stored and are a true reflection of performance at the workstation.”*

Lee and Weerakoon (2001, category 1a) stress, “It is important to remember that the ultimate strength of the multiple-choice assessment as a reliable and valid tool for educational assessment lies in the quality of the test items. No computer software, however sophisticated or easy to use, will compensate for poorly written test items.” This point holds true for all assessments and feedback, whether technology-enhanced or otherwise, and regardless of the technology used.

Testing the assessment and feedback to ensure it is reliable and valid and that the technology delivers as expected is therefore an important stage in the development process.

## 5. Piloting

### 5.1 Key messages about the piloting stage

By ‘piloting’ we mean testing the assessment and feedback with students and academic colleagues to gauge what the responses and reactions of the targeted students will be and to ensure the assessment will work as intended for them. This is an important stage in the quality assurance of the assessment and feedback. Using technology can offer advantages

to tutors during this stage because of the possibilities of student response analysis (Conole and Warburton, 2005). This paper traces the development of the use of technology-enhanced methods based on a literature review (evidence category 3).

*“One of the benefits of CAA [computer-assisted assessment] is the opportunity to record student interactions and analyse these to provide a richer understanding of learning. A variety of analyses can be run to assess how well individual questions or students perform. Weak items can then be eliminated or teacher strategies adapted.”*

They point out that data on student performance can be recorded automatically but that care is needed in interpreting it, using the example: “incorrect responses may indicate a more sophisticated understanding by the student than might at first appear; for example, incorrect use of grammar in a foreign language test might be a result of higher cognitive understanding by the student.”

McKenna and Bull (2000) also highlight the benefits of monitoring question performance when using technology-enhanced methods. Their paper is in evidence category 3. In it they give the example:

*“Statistical analyses of results can determine discrimination, facility and the level of correlation within tests and to other question types. Discrimination measures the power of an individual question to discriminate between the stronger and weaker students. It does so by establishing the correlation between the performance of a subgroup of students on a particular question and their overall scores on the assessment. Facility measures the difficulty of a question by dividing the number of correct responses by the total number of responses. Questions with poor facility or discrimination can be improved or eliminated.”*

They recommend undertaking question reviews during pre-testing as well as after using an assessment.

## **6. Delivery**

### **6.1 Focus on practical issues**

In this section we focus on the practical issues of delivering technology-enhanced assessment and feedback such as via computers or using podcasting.

Beevers et al (1995, evidence category 3) identify the three main issues they found of using computers to deliver maths exams:

1. *“Screens visible to neighbouring students present a security problem;*
2. *How computer marking could cope with partial credit; and*
3. *The mis-match between mathematical and computer notations.”*

The team tackled the first problem by using random-selection from eight question banks, making the chances of neighbouring students receiving the same question at the same time

very low. To ensure all students received a test at the same level of difficulty, two experienced teachers “strove to attain a similar standard” across the question banks.

Other practical issues are addressed by Gilbert et al (2009, evidence category 4). “When it came to delivery, interviewees mentioned the following practical advice:

- *“Stagger the start of online exams – need to make sure invigilators know about this – to avoid overloading server.”*
- *“Sometimes people are told the server can easily cope with 100 or 200 logons at the same time, unaware that the link between their students and the server is not as good. So, it's not just the server, it is also the connectivity that needs to be considered.”*
- *“Have paper copies of the exam as a contingency – about 10% of the number of students.”*
- *“Piloting and ‘iterate and refine’ is important, but it's crucial to have a good reporting system in order to see and analyse the results (including student answers) in order to complete this process.”*

### **How one university managed practical issues**

The booklet *Effective Practice with E-Assessment* (JISC, 2007 evidence category 4) has a case study from the University of Manchester which explains how they managed the issues of delivering exams by computer:

*“Student identity was authenticated by students logging on to the VLE using their normal university username and password, and was backed up by the invigilator checking that students taking the examination could be identified from photos on their university ID cards. Timed release was used to ensure that examinations could be accessed only during the timetabled period and, for additional security, a test specific password was issued. This was given to candidates only when the examination commenced.*

*The possibility of a technical failure on the day of the examination remained a concern. To prepare for this eventuality, paper versions of the tests were produced as a backup and, as a further fail-safe mechanism, candidates were asked to enter their responses online and on an optical mark reader sheet. This prevented the examination from being a fully computer-based one, in which random ordering of questions could take place, but nonetheless enabled a useful comparison between different methodologies during the pilot. The process relied on the assistance of a support team, a backup server operated throughout the examination and computers were booked at 90% capacity to allow for the breakdown of more than one machine.”*

Further details are given by Shepherd et al (2006) who implemented technology-enhanced assessment in six departments within the University of Southampton. This paper is in evidence category 3 and looks at the implementation in relation to the relevant quality assurance standard, BS7988. The authors note the “the possibility of system failure at any stage of the assessment should be acknowledged, planned for and provided for in all programmes that make use of CAA, particularly for summative assessment.” They also identify how important it is to:

*“Define or reach a common view on the distribution of responsibility for particular and defined aspects of CAA. Some questions are as follows:*

- *To what extent will tutors responsible for an academic assessment be expected to be familiar with CAA software and associated delivery systems?*
- *To what extent will or should aspects of this responsibility be delegated to other members of a programme team or indeed to other support staff, for example, CAA specialists in Information Systems Services?*
- *To what extent will tutors be expected to be familiar with the pedagogical theory underlying the construction of CAA?”*

The authors also recommend defining a process that will “control the events between the preparation of an assessment and its operation”, and recommend preparing students to take the assessments that use technology enhancement, similar to being able to practice on past papers.

Advice is therefore available from experienced practitioners about the pitfalls involved with using technology enhancement and the precautions they take to manage them.

## **7. Capturing Student Responses**

Capturing student responses is an important step in the process of using technology to enhance assessment and feedback. However, the references highlighted to us did not address this step specifically. Instead, the decision to use technology to capture student responses, portfolios of evidence or group work or to enable peer feedback is seen as part of the pedagogical design, discussed in Section 2 Design.

## **8. Feedback to Students**

The principles and theory of designing feedback are discussed in Section 2 Design. Practical issues in the actual delivery of feedback have not received specific study, and tend to be incorporated into papers which deal more generally with assessment delivery, as discussed in Section 6 Delivery.

## **9. Evaluation**

### **9.1 Measuring validity and reliability**

The experienced practitioners interviewed for the *Report on E-Assessment Quality* project (Gilbert et al, 2009, evidence category 4), identified validity and reliability as the key criteria for evaluating any assessment with technology enhancement. These criteria include the ‘fairness in item banking’ that is of particular concern to students. The authors identify statistical measures for reliability, i.e., consistency, coherence, and robustness that answer the following questions:

- For the test as a whole:
  - Does it give the same result when different versions are used?
  - Are the same results obtained when the test is re-taken?
  - Do the 'better' students do better on the test?
- For individual question items in a test:
  - Do they show an appropriate range of difficulty?
  - Is the selection of incorrect item options appropriately distributed?
  - Do the 'better' students tend to answer correctly?

For validity the questions that need to be addressed are:

- Do the test items follow directly from the intended learning outcomes?
- Do the test results predict future performance on similar tests?
- Do the test results correlate with other measures of student knowledge and ability?

## 9.2 Evaluating reports generated by the assessment technology

Cronbach (1971), quoted in Wiliam (2008, evidence category 4), noted, "One validates, not a test, but an *interpretation of data arising from a specified procedure*" (p447, emphasis in original).

McKenna and Bull (2000, evidence category 3) explain that:

*"Most CAA software will automatically generate a range of reports on question and student performance. If a regular programme of CAA is embedded within the delivery of a module, weaknesses in teaching or student learning can be detected swiftly, thus allowing for the implementation of timely modifications and potentially contributing to a process of integrative evaluation of the module."*

## 9.3 The importance of regular reviews

Zakrewski and Steven (2003, evidence category 3) recommend an annual review of technology-enhanced assessment systems:

*"A review board should be set up that would be responsible for the compilation of an annual monitoring report. This would coincide with the institutions annual monitoring cycle. The report will recommend changes and will be the focus for the planning stage in the next cycle of the 'Catherine Wheel' CBA [computer-based assessment, Section 10 Feedback to Stakeholders] model. It will be received by the academic standards committee of the institution for detailed consideration. The annual monitoring report considers the quality assurance issues together with a student evaluation of the system, external examiners reports, a comparative evaluation of student achievement in CBA with more traditional forms of assessment, recommendations for improvements and an action plan for next academic year."*



## 9.4 Summary of key messages on the evaluation stage

The evaluation stage focuses on issues of assessment reliability and validity. The reports that can be generated by technology can be very helpful in checking that each question and the test as a whole is reliable and valid, and refining any items that do not meet these standards. Just as with other forms of assessment, regular reviews are needed to ensure assessments with technology enhancement are maintained.

## 10 Feedback to Stakeholders

### 10.1 Introduction

In this section we will explore who the different stakeholder groups are and what their specific interests are in technology-enhanced assessment. We will also focus on quality assurance measures as these are of interest to all groups.

### 10.2 Addressing the different interests of different stakeholder groups

The implications of introducing technology-enhanced assessment for stakeholders and therefore the issues they are particularly interested in are summarised by Zakrewski and Steven (2003, evidence category 3). They identified the different interests of students, academic staff, support staff and managers. The summary is provided in full in Appendix E.

Warburton (2009, evidence category 2) identified three stakeholder groups and what their “notions of successful uptake” of technology-enhanced assessment are:

- Tutors: “completing assessment tasks with maximal efficiency (and minimal student unrest)”;
- Learning technologists: “pedagogic fitness for purpose and extending technical boundaries”;
- Senior managers and quality assurance specialists: “tended to be principally concerned for their institutions’ reputations and saw CAA as a risk to be managed by the centre”.

To address the needs of all three stakeholder groups, Stephens et al (1998, evidence category 3) suggested guidelines for an institutional computer-assisted assessment (CAA) strategy based on reviewing case studies:

1. *“Establish a coordinated CAA management policy for the CAA unit(s) and each discipline on campus.*
2. *Establish a CAA unit: a central CAA unit for the institution or distributed CAA units within faculties/schools can provide a focal point and centre of expertise.*
3. *Appoint discipline coordinators within departments to provide discipline-specific advice and guidance to academic staff and, where relevant, the CAA unit(s).*

4. *Establish CAA discipline groups/committees to link schools/faculties, their units and the central unit.*
5. *Provide funding for CAA Units.*
6. *Organise staff development programmes on the identification and development of good practice and design methodologies, as well as operational procedures and help to ensure staff are aware of technical capabilities and limitations of software.*
7. *Establish evaluation procedures.*
8. *Identify technical issues.*
9. *Establish operational and administrative procedures.”*

### **10.3 Ensuring accessibility**

JISC TechDis has produced a booklet on the *12 steps towards embedding inclusive practice with technology as a whole institution culture in UK Higher Education*, which is available from its website: [www.techdis.ac.uk](http://www.techdis.ac.uk).

### **10.4 Ensuring high quality**

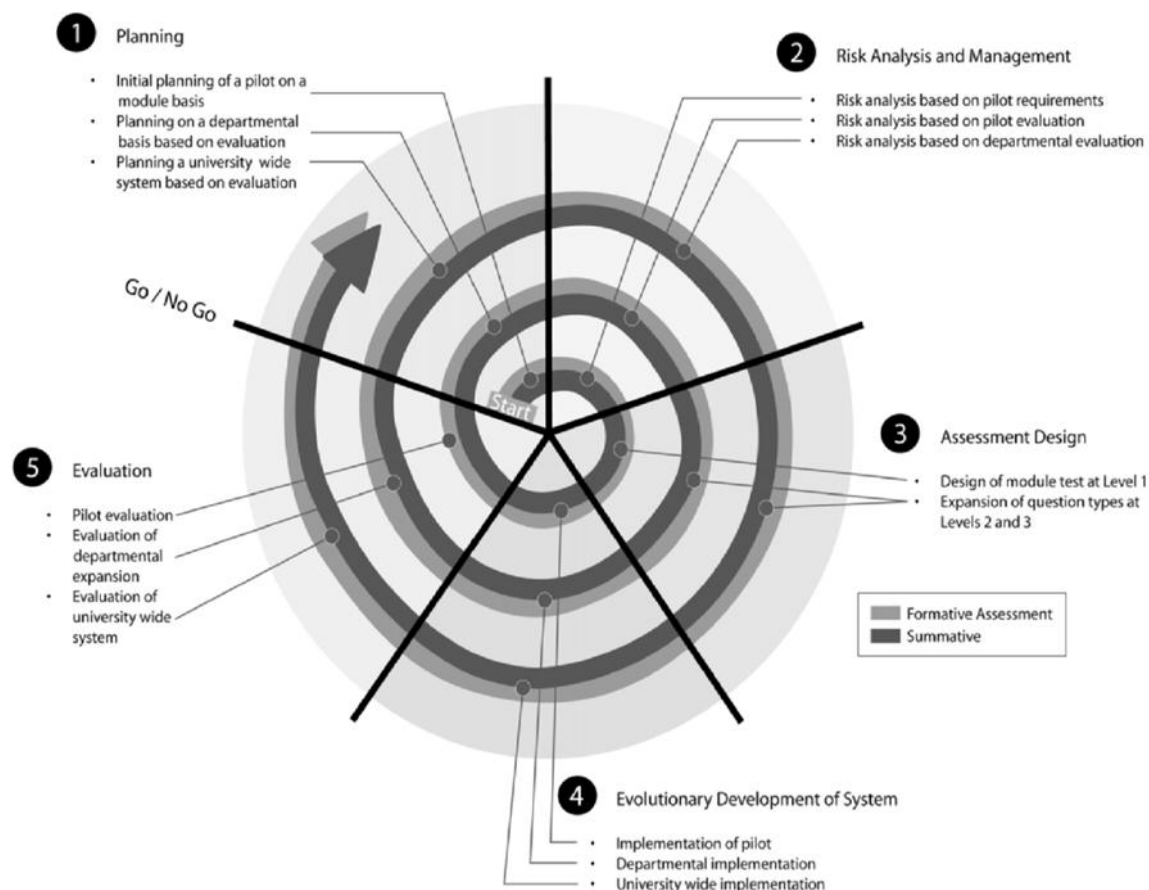
Quality assurance issues in computer-based assessment are discussed by Zakrzewski and Steven (2000 and 2003, evidence category 3). They summarise the findings from the literature on quality assurance issues for stakeholders in technology-enhanced assessment as:

- *“Integration with current university assessment systems.*
- *Staffing issues, particularly staff development in objective test design.*
- *Examination design and appropriateness of the objective test questions at levels 1, 2 and 3 [first year, second year and third year of an undergraduate programme] to the knowledge and skills being assessed.*
- *Endorsement by external examiners and advisers on examination design.*
- *The thoroughness of the system testing procedures before the examination period.*
- *System integrity in terms of student conduct, invigilation, security and back-up procedures.*
- *Student issues and in particular student anxiety about a new assessment method that will require a dedicated user interface appropriate to CBA implementation.*
- *Annual review of procedures, student performance and costs underpinned by feedback from students, academic staff, support staff and externals.”*

The authors discuss each of these issues, highlighting the need for all stakeholders to know how the university will operate technology based systems; staff development and support for academic staff; the “vital” role of question analysis software and statistical reporting software and the “problematical issue” of exam updates.

In response to these issues, they have implemented a ‘Catherine Wheel’ approach at the University of Luton. The Catherine Wheel model takes a “structured approach” to the quality of computer-based assessment. It describes “five segments: planning, risk analysis and

management, assessment design, evolutionary development, and evaluation.” This process can be applied to all technology-enhanced assessment, regardless of the technology used.



**Figure 6**

*Quality Assurance 'Catherine Wheel' from Zakrewski and Steven (2003)*

## 10.5 Summary of key messages on feedback to stakeholders

In this section we looked at the recommended literature to identify some of the stakeholder groups involved in assessment and feedback and how to handle their needs at an organisational level, from the point of view of accessibility and quality assurance.

### **III. Conclusions**

The references that were recommended to us are clearly having an impact on current practice and are found valuable by practitioners. The key messages from these sources are consistent and often give detailed and practical guidance for other academics. We found that most of the recommended literature focused on the goals that technology enhancement can enable assessment and feedback to meet and how assessment and feedback can be designed to make best use of the technology.

#### **A. Limitations**

The literature on technology-enhanced assessment and feedback is very large. One limitation of this report is that it was not possible to review all possible sources for evidence. Instead we were guided by and limited to the recommendations of our project participants.

#### **B. Observations**

- Most recommended publications lack quantified evidence.

While all the recommended publications may be characterised as reputable (Table 2) and the majority were peer-reviewed (67.7%), only a minority provided quantitative data (28.2%), of which relatively few provided appropriate experimental designs or statistical analysis (18.5%). The great majority of publications proposed to us were practitioner-led case studies. Few of the publications were based on evidence categories 1a, 1b and 2 (Table 1).

- Most evidence addresses reaction rather than effective change.

We found Kirkpatrick's Evaluation Model<sup>1</sup> (Appendix G) to be particularly appropriate, and note that the great majority of the recommendations we received provided evidence at Kirkpatrick Level 1, i.e., surveys of students' and academics' reaction (likes, dislikes, and perceived effectiveness) rather than demonstrating changes in learning, changes in performance, or changes in organisational or stakeholder benefits.

- Most recommended publications lack direct application.

The majority of publications proposed to us were case studies which were highly contextualised to the specifics of a course or organisation and any generalisations were usually at an exceptionally high level.

- It's the pedagogy which makes the difference.

We concur with those publications which emphasised the significance of the pedagogy and the secondary importance of the technology.

#### **C. Comments**

On the basis of these observations, we provide two comments.

---

<sup>1</sup> [www.kirkpatrickpartners.com](http://www.kirkpatrickpartners.com)

- Introduce risk and change management strategies when committing significant resources.

Because the majority of references are largely case studies and lack quantified outcomes, prudence may be needed when changing to technology-enhanced assessment strategies at course or institutional level.

- Address learning first and the affordances of technology enhancement second.

In engaging with technology-enhanced learning, the majority of references concluded that success followed from focusing on learning.

## **D. Further work**

It would be interesting to find out why these articles have more impact than others.

## **E. Key messages**

Some of the key messages that were recommended by practitioners to their peers are:

- Assessment for learning shows an effect size of between 0.34 and 0.46 for assessment for learning (Black 2003), see page 8;
- Tutors can use technology-enhanced methods to implement effective learning designs that would not otherwise be possible because of factors such as time constraints, student numbers and geographical or temporal distribution, see page 9;
- Effective regular, online testing can encourage student learning and improve their performance in tests, see page 10;
- Student retention and inclusion can be increased by using technology-enhanced methods. Exam anxiety can also be reduced, see page 17;
- Using technology-based methods does not disadvantage women or older students, see page 21;
- Automated marking can be more reliable than human markers and there is no medium effect between paper and computerized exams, see pages 16, 22 and 23;
- The success of assessment and feedback with technology enhancement lies with the pedagogy rather than the technology itself; technology is an enabler, see page 26;
- Technology-enhanced assessment is not restricted to simple questions and clear-cut right and wrong answers, much more sophisticated questions are being used as well, see page 28;
- Modern technology can be matched to the learning characteristics of the contemporary learner, see page 28;
- The design of appropriate and constructive feedback plays a vital role in the success of assessment, especially assessment for learning. The literature offers detailed guidance on designing effective feedback such as conditions, research-backed principles and a typology, as well as specific advice for the design of audio feedback and peer assessment, see page 33;
- The characteristics of useful technologies to use for assessment and feedback, see page 40;
- The value of taking a team approach to the creation of technology-enhanced assessment and feedback as successful implementation requires skills in the

application of technology and how to use to the technology itself as well as learning and the subject content, see page 42;

- The vital role of staff development and support when introducing and developing assessment and feedback with technology enhancement, see page 43;
- Testing the assessment and feedback to ensure it is reliable and valid and piloting it with people who are similar to or understand the targeted students are important stages in the development process. A good reporting system can help academics see and analyse the results (including student answers) and will help refine the assessment and feedback, see page 45;
- It is important to prepare students to take the assessments that use technology enhancement by practicing with similar levels of assessment using the same equipment and methods. This is similar to being able to practise on past papers, see page 48;
- The reports generated by many technology-enhanced assessment systems are very helpful in checking the reliability and validity of each test item and the test as a whole, see page 49.

## IV. Appendices

### A. Survey questions

Our surveys explained that that we were particularly interested in **evidence-based** literature on assessment and feedback.

Our survey posed the following questions:

Q1: What references about assessment and feedback would you recommend to other practitioners?

Q2: What assessment and feedback topics do you recommend we cover in our synthesis report to ensure it is as useful as possible?

Q3: What format would you prefer the synthesis output to take?

### B. Glossary

#### Assessment

Quality Assurance Agency (2006): Assessment is “any process that appraises an individual’s knowledge, understanding, abilities or skills.”

From JISC (2007) *Effective Practice with E-Assessment*:

Formative – assessment that provides developmental feedback to a learner on his or her current understanding and skills. Formative assessment can also be described as ‘assessment for learning’ since an assessment that is entered into voluntarily, and on which no final qualification depends, can prompt learners to adjust their own performance.

Summative – the final assessment of a learner’s achievement, usually leading to a formal qualification or certification of a skill. Summative assessment is also referred to as assessment of learning.

#### Low, medium or high stakes

Assessment of any kind can be referred to as low, medium or high stakes. A low-stakes assessment is usually formative, with results recorded locally. A medium-stakes assessment is one in which results may be recorded locally and nationally, but is not life changing. A high-stakes assessment, however, is one in which the outcomes are of high importance to both centre and candidates, affecting progression to subsequent roles and activities.

## **e-Portfolio**

The e-portfolio can be used to document in digital form the outcomes of each stage in a learner's journey. The distinction between an e-portfolio as a record of achievement and an e-portfolio as a tool for assessment becomes blurred when the outcomes of assessments, including self- or peer-assessments in the form of diaries, blogs or wikis, are included. An e-portfolio may also be the means by which some qualifications are assessed.

## **C. Lists of principles and conditions**

### **Assessment Reform Group (2002): Assessment for Learning 10 principles**

Assessment Reform Group's research-based principles to guide classroom practice (retrieved from [www.assessment-reform-group.org](http://www.assessment-reform-group.org)) suggest assessment for learning should:

1. Be part of effective planning of teaching and learning.
2. Focus on how students learn.
3. Be recognized as central to classroom practice.
4. Be regarded as a key professional skill for teachers.
5. Be sensitive and constructive because any assessment has an emotional impact.
6. Take account of learner motivation.
7. Promote commitment to learning goals and a shared understanding of the criteria by which they are assessed.
8. Help learners know how to improve through constructive guidance.
9. Develop learner's capacity for self-assessment so that they can become reflective and self-managing.
10. Recognise the full range of achievements of all learners.

### **Chickering and Gamson's (1987) Seven Principles of Good Practice in Undergraduate Education**

1. Encourages contact between students and faculty.
2. Develops reciprocity and cooperation among students.
3. Encourages active learning.
4. Gives prompt feedback.
5. Emphasises time on task.
6. Communicates high expectations.
7. Respects diverse talents and ways of learning.



## **Seven principles of good feedback practice: facilitating self-regulation, Nicol and MacFarlane-Dick (2006)**

Good feedback practice:

1. Helps clarify what good performance is (goals, criteria, expected standards);
2. Facilitates the development of self-assessment (reflection) in learning;
3. Delivers high quality information to students about their learning;
4. Encourages teacher and peer dialogue around learning;
5. Encourages positive motivational beliefs and self-esteem;
6. Provides opportunities to close the gap between current and desired performance;
7. Provides information to teachers that can be used to help shape teaching.

## **Gibbs and Simpson's (2004) 11 assessment conditions**

### **Assessment tasks (conditions 1–4)**

- (1) Capture sufficient study time and effort (in and out of class)
- (2) Are spread evenly across topics and weeks
- (3) Lead to productive learning activity (deep rather than surface learning)
- (4) Communicate clear and high expectations

### **Feedback (conditions 5–11)**

- (5) Is sufficient (in frequency, detail)
- (6) Is provided quickly enough to be useful
- (7) Focuses on learning rather than marks
- (8) Is linked to assessment criteria/expected learning outcomes
- (9) Makes sense to students
- (10) Is received by students and attended to
- (11) Is acted upon to improve work and learning

## **D. Academy's Seminar Series links**

### **Technology, Feedback, Action! Sheffield Hallam University**

Seminar briefing paper on the Academy's website:

[http://www.heacademy.ac.uk/assets/York/documents/ourwork/learningandtech/Sheffield\\_Hallam\\_University\\_17\\_Feb\\_2010\\_briefing\\_report.pdf](http://www.heacademy.ac.uk/assets/York/documents/ourwork/learningandtech/Sheffield_Hallam_University_17_Feb_2010_briefing_report.pdf)

Guides available from the Higher Education Academy EvidenceNet:

*Feedback: a student guide to using feedback -*

[http://www.heacademy.ac.uk/assets/EvidenceNet/TFA\\_student\\_guide.pdf](http://www.heacademy.ac.uk/assets/EvidenceNet/TFA_student_guide.pdf)

*Using technology to help students engage with their feedback: A best practice guide for academic staff*

[http://www.heacademy.ac.uk/assets/EvidenceNet/TFA\\_guide\\_for\\_academic\\_staff.pdf](http://www.heacademy.ac.uk/assets/EvidenceNet/TFA_guide_for_academic_staff.pdf)

*Using technology to help students engage with their feedback: A ten minute guide for senior managers*

[http://www.heacademy.ac.uk/assets/EvidenceNet/TFA\\_guide\\_for\\_senior\\_managers.pdf](http://www.heacademy.ac.uk/assets/EvidenceNet/TFA_guide_for_senior_managers.pdf)

### **Podcasting in Assessment: New Technology in Higher Education Research (PANTHER) University of Leicester**

Seminar briefing paper on the Academy's website:

[http://www.heacademy.ac.uk/assets/York/documents/ourwork/learningandtech/Leicester\\_Briefing\\_Report\\_3\\_March\\_2010.pdf](http://www.heacademy.ac.uk/assets/York/documents/ourwork/learningandtech/Leicester_Briefing_Report_3_March_2010.pdf)

Parent project website: <http://www.le.ac.uk/duckling>

Moving Forward Through Feedback: Enhancing Feedback Provision in the Digital Age, University of Reading

Seminar briefing paper:

[http://www.heacademy.ac.uk/assets/York/documents/ourwork/learningandtech/University\\_of\\_Reading\\_Briefing\\_Report\\_14\\_April\\_2010.pdf](http://www.heacademy.ac.uk/assets/York/documents/ourwork/learningandtech/University_of_Reading_Briefing_Report_14_April_2010.pdf)

The remaining briefing papers are available on the Academy's website at:

<http://evidencenet.pbworks.com/Evidence-based+Practice+Seminar+Series+2010%3A+Assessment+and+Feedback>

## **E. Zakrewski and Steven (2003): interests of different stakeholder groups**

### **Students**

- Changing the way they learn
- Further enhancement of IT skills
- Extensive reflection on performance
- Avoidance of embarrassment
- Working at a place and time appropriate for the individual

- Enjoyment

### **Academic staff**

- Time to develop effective tests
- Fast feedback on performance to allow problem areas to be dealt with quickly
- Reduction in marking loads thus freeing the academic to develop other areas of interest
- Balance of assessment strategies to test the range of skill levels

### **Support staff**

- Implementation and delivery of CBA [computer-based assessment] systems
- Time taken for testing system reliability
- In-depth knowledge of the software used

### **Managers**

- Coordination and management of team approach
- Appointment of additional specialist staff
- Establishment of a central unit for CBA
- Investment in hardware and software
- Space investment to support delivery of CBA to large cohorts
- Integration with existing assessment procedures
- Availability of appropriate documents to include new quality assurance procedures”

## **F. Cost-benefit analysis for e-learning**

The benefits of e-learning can be categorised into one or more of these general areas (Gilbert & Gale, 2007):

- cost-savings due to reducing staff or other required resources.
- productivity improvements due to enabling higher levels of production or service without a commensurate increase in resources.
- quality of information or provision, involving better decision making, more accurate and more timely information, the reduction of unnecessary communications, and the improvement of required communications.
- organisational enhancement, involving the ability to undertake new business ventures, penetrate new markets, and develop new products.

## G. Kirkpatrick levels of evaluation

A useful classification of the kind of information to seek from an evaluation is associated with the work of Kirkpatrick's Evaluation Model<sup>2</sup>. Gilbert & Gale (2007) provide the following synopsis:

- Level 1 - Reaction
- Level 2 - Learning
- Level 3 - Performance change
- Level 4 - Organisational benefit

**Reaction** information is typically the information gathered by course evaluation questionnaires – essentially questions around the issue of how much the students liked the e-learning. This information is readily obtained through questionnaires and is readily analysed in quantitative and statistical terms.

**Learning** information concerns the effectiveness of the e-learning – questions around the issues of how much, and how well, the students actually learned what was intended, and possibly around the issue of how much they learned that was unintended.

**Performance change** information is concerned with whether the learning which took place as a result of the educational or training experience due to the e-learning resulted in actual performance changes and improvements on the part of the students.

**Organisational benefit** information concerns the overall value of the e-learning for both the organisation and the stakeholders.

## H. Evidence category 1a references

The following recommended references fall into evidence category 1a, that is, peer-reviewed generalizable studies providing effect size estimates and which include (i) some form of control group or treatment (may involve participants acting as their own control, such as before and after), and / or (ii) blind or preferably double-blind protocol.

Angus, S. D. and Watson, J. (2009). Does regular online testing enhance student learning in the numerical sciences? Robust evidence from a large data set. *British Journal of Educational Technology*, 40(2), 255-272.

Ashton, H. S., Beevers, C. E., Korabinski, A. A. and Youngson M. A. (2005). Investigating the medium effect in computer-aided assessment of school Chemistry and college computing national examinations. *British Journal of Educational Technology Vol 36 No 5* 771–787.

Ashton, H. S., Beevers, C. E., Korabinski, A. A., and Youngson, M. A., (2006). Incorporating partial credit in computer-aided assessment of Mathematics in secondary education. *British Journal of Educational Technology Vol 37 No 1* 93–119.

---

<sup>2</sup> [www.kirkpatrickpartners.com](http://www.kirkpatrickpartners.com)

- Barbera, E (2009) Mutual feedback in e-portfolio assessment: an approach to the netfolio system. *British Journal of Educational Technology*, 40(2), 342-357.
- Black, P., Harrison, C., Lee, C. Marshall, B. and William, D. (2003). *Assessment for Learning. Putting it into practice*. Open University Press.
- Butcher, P.G & Jordan, S.E. (2010) A comparison of human and computer marking of short free-text student responses. *Computers and Education*, 55, 489-499.
- Chang, C. and Tseng, H. (2009). Use and performances of Web-based portfolio assessment. *British Journal of Educational Technology Vol 40 No 2* 358–370.
- Hake, R. R. (1997). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *The American Journal of Physiology*, 66(1) 64-74.
- Jordan, S. and Mitchell, T. (2009). e-assessment for learning? The potential of short-answer free-text questions with tailored feedback. *British Journal of Educational Technology*, 40(2), 371-385.
- Kochakornjarupong, D. and Brna, P. (2010) Helping feedback-givers to improve their feedback. In D.Whitelock and P.Brna (eds) Special Issue 'Focusing on electronic feedback: feasible progress or just unfulfilled promises?' *Int. J. Continuing Engineering Education and Life-Long Learning*, Vol. 20, No. 2, pp.148-168.
- Lee, G. and Weerakoon, P. (2001). The role of computer-aided assessment in health professional education: a comparison of student performance in computer-based and paper-and-pen multiple-choice tests. *Medical Teacher*, Vol 23, No. 2, 152 - 157.
- Loddington, S., Pond, K., Wilkinson, N., & Willmot, P. (2009). A case study of the development of WebPA: An online peer-moderated marking tool. *British Journal of Educational Technology*, 40(2), 329-341.
- Mitchell, T., Aldridge, N., Williamson, W. and Broomhead, P. (2003). Computer based testing of medical knowledge. 7th International CAA Conference, Loughborough. Last accessed September 2010 from <http://www.caaconference.com/pastconferences/2003/index.asp>.
- Nicol, D. (2009) Assessment for learner self-regulation: enhancing achievement in the first year using learning technologies. *Assessment and Evaluation in Higher Education*, 34: 3, 335 — 352.
- Sandene, B., Horkay, N., Bennett, R. E., Allen, N., Braswell, J., Kaplan, B. and Oranje, A. (2005). Online Assessment in Mathematics and Writing: Reports from the NAEP Technology-Based Assessment Project, Research and Development Series. Last accessed August 2010 from <http://nces.ed.gov/nationalreportcard/pubs/studies/2005457.asp>.
- Strang, K. D. (2010). Measuring self regulated e-feedback, study approach and academic outcome of multicultural university students. In D.Whitelock and P.Brna (eds) Special Issue 'Focusing on electronic feedback: feasible progress or just unfulfilled promises?' *Int. J. Continuing Engineering Education and Life-Long Learning*, Vol. 20, No. 2, pp.239-255.

## V. References

Angus, S. D. and Watson, J. (2009). Does regular online testing enhance student learning in the numerical sciences? Robust evidence from a large data set. *British Journal of Educational Technology*, 40(2), 255-272.

Ashton, H. (2008). Exploring e-Assessment: A practitioner's perspective. JISC, e-Learning Focus, October 2008 <http://www.elearning.ac.uk/features/eassessment>.

Ashton, H. S., Beevers, C. E., Korabinski, A. A. and Youngson M. A. (2005). Investigating the medium effect in computer-aided assessment of school Chemistry and college computing national examinations. *British Journal of Educational Technology Vol 36 No 5* 771–787.

Ashton, H. S., Beevers, C. E., Korabinski, A. A., and Youngson, M. A., (2006). Incorporating partial credit in computer-aided assessment of Mathematics in secondary education. *British Journal of Educational Technology Vol 37 No 1* 93–119.

Ashton, H.S., Beevers, C.E., Schofield, D.K. & Youngson, M.A. (2004). Informative Reports – experiences from the PASS-IT Project. Proceedings of the 8th International CAA Conference, Loughborough University <http://www.caaconference.com/pastConferences/2004/proceedings/ashton.pdf>.

Ashton, H. and Thomas, R. (2006) Bridging the gap between assessment, learning and teaching. Proceedings of the 10<sup>th</sup> International CAA Conference, Loughborough: Loughborough University. <http://www.caaconference.com/pastConferences/2006/proceedings/ashton.pdf>.

Ashton, H.S. and Youngson, M.A. (2004). Creating Questions for Automatic Assessment in Mathematics. LTSN MSOR Maths CAA series, February article. Retrieved from [http://mathstore.ac.uk/repository/mathscaa\\_feb2004.pdf](http://mathstore.ac.uk/repository/mathscaa_feb2004.pdf).

Ashton, H. S., and Wood, C. (2006) Use of Online Assessment to Enhance Teaching and Learning: the Pass-IT Project. *European Educational Research Journal*, 5(2), 122-130 [http://www.wwwords.co.uk/EERJ/content/pdfs/5/issue5\\_2.asp#5](http://www.wwwords.co.uk/EERJ/content/pdfs/5/issue5_2.asp#5).

Assessment Reform Group (2002). Assessment for learning: 10 principles. Research-based principles to guide classroom practice. Downloaded from [www.assessment-reform-group.org](http://www.assessment-reform-group.org).

Barbera, E (2009) Mutual feedback in e-portfolio assessment: an approach to the netfolio system. *British Journal of Educational Technology*, 40(2), 342-357.

Beaumont, C., O'Doherty, M., and Shannon, L. (2008). Staff and student perceptions of feedback quality in the context of widening participation, Higher Education Academy. Retrieved May 2010 from: [http://www.heacademy.ac.uk/assets/York/documents/ourwork/research/Beaumont\\_Final\\_Report.pdf](http://www.heacademy.ac.uk/assets/York/documents/ourwork/research/Beaumont_Final_Report.pdf).

- Beevers, C., McGuire, G., Stirling, G., and Wild, D., (1995). Mathematical Ability Assessed by Computer. *Journal of Computers and Education* 25, pp.123-132.
- Black, P., Harrison, C., Lee, C. Marshall, B. and Wiliam, D. (2003). *Assessment for Learning. Putting it into practice*. Open University Press.
- Bloxham, S. and Boyd, P. (2008) *Developing Effective Assessment in Higher Education. A Practical Guide*. Open University Press.
- Boud, D. (2000) Sustainable Assessment: Rethinking assessment for the learning society, *Studies in Continuing Education*, 22: 2, 151 — 167.
- Boud, D. and Falchikov, N. (editors) (2007). *Rethinking Assessment in Higher Education. Learning for the longer term*. London. Routledge.
- Boyle, A. (2007). The formative use of e-assessment: some early implementations, and suggestions for how we might move on. Proceedings of 11<sup>th</sup> International CAA Conference, Loughborough. Retrieved from <http://www.caaconference.com/pastConferences/2007/proceedings/index.html>.
- Boyle, A. and Hutchison, D. (2009). Sophisticated tasks in e-assessment: what are they and what are their benefits?, *Assessment & Evaluation in Higher Education*, 34: 3, 305 — 319.
- Bull, J. (1999). Computer-assisted assessment: impact on higher education institutions. *Educational Technology & Society*, 2(3), 123-126.
- Butcher, P. G. and Jordan, S. E. (2010). A comparison of human and computer marking of short free-text student responses. *Computers and Education*, 55, 489-499.
- CAA Centre Resources Guide to objective tests. Available from [http://www.caacentre.ac.uk/resources/objective\\_tests/](http://www.caacentre.ac.uk/resources/objective_tests/).
- Cassady, J. C. and Gridley, B. E. (2005). The effects of online formative and summative assessment on test anxiety and performance. *Journal of Technology, Learning and Assessment*, 4(1), 1-31.
- Challis, D. (2005). Committing to quality learning through adaptive online assessment. *Assessment & Evaluation in Higher Education*, 30: 5, 519 — 527.
- Chang, C. and Tseng, H. (2009). Use and performances of Web-based portfolio assessment. *British Journal of Educational Technology Vol 40 No 2* 358–370.
- Chickering, A. W. and Gamson, Z. F. (1987). Seven principles for good practice in undergraduate education. *The American Association for Higher Education Bulletin*, March.
- Conole, G. and Warburton, B. (2005). A review of computer-assisted assessment. *ALT-J*, 13: 1, 17 — 31.
- Crouch, C. H. and Mazur, E. (2001). Peer Instruction: Ten years of experience and results. *American Journal of Physics*, 69, 970-977.
- Dermo, J. (2009). e-Assessment and the Student learning experience: A survey of student perceptions of e-assessment. *British Journal of Educational Technology*, 40 (2), 203 -214.

Dexter, S. (2010). E-feedback intersections and disconnections in the interests of designers and users. In D. Whitelock and P. Brna (eds) Special Issue 'Focusing on electronic feedback: feasible progress or just unfulfilled promises?' *Int. J. Continuing Engineering Education and Life-Long Learning*, Vol. 20, No. 2, pp. 169-188.

Draper, S. (2009). Catalytic assessment: understanding how MCQs and EVS can foster deep learning. *British Journal of Educational Technology*, 40(2), 285-293.

Elliott, B (2007). Guide to Writing Objective Tests. Scottish Qualifications Authority. Available from <http://www.scribd.com/doc/464785/Objective-Test-Guide>.

Elliott, B. (2008). Assessment 2.0: Modernising assessment in the age of Web 2.0. Scottish Qualifications Authority. Retrieved June 18, 2010, from <http://www.scribd.com/doc/461041/Assessment-20>.

Entwistle, N.J., McCune, V. and Walker, P. (2004). Conceptions, styles and approaches within higher education: analytic abstractions and everyday experience. In R. J. Sternberg and L-F Zhang (eds) *Perspectives on Thinking, Learning and Cognitive Styles*, Erlbaum, London.

Gardner, J. (ed) (2006). *Assessment and Learning*. Sage.

Gibbs, G., and Simpson, C. (2004). Conditions under which assessment supports students' learning. *Learning and Teaching in Higher Education 1*: 3–31.

Gilbert, L., and Gale, V. (2007). *Principles of E-Learning Systems Engineering*. Chandos Publishing (Oxford) Ltd.

Gilbert, L., Gale, V., Warburton, B. and Wills, G. (2009). Report on Summative E-Assessment Quality (REAQ). JISC. Last accessed September 2010 from <http://www.jisc.ac.uk/whatwedo/projects/reaq.aspx>.

Gipps, C.V. (2005). What is the role for ICT-based assessment in universities? *Studies in Higher Education*, 30(2), 171-180.

Gordon, N. A., (2008). Group work in Higher Education Computing Courses. Proceedings of the HEA ICS Annual Conference 2008, Edited by H. White, ISBN 978-0-9559676-0-3, HE Academy, Subject Centre for ICS, pp35-39.

Hake, R. R. (1997). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *The American Journal of Physics*, 66(1) 64-74.

Hattie J. A. (1987). Identifying the salient facets of a model of student learning: A synthesis of meta-analyses. *International Journal of Educational Research*, Vol. 11 (2) pp.187 – 212.

Hofstede, G. (1991). *Cultures and Organisations: Software of the Mind*, McGraw-Hill, London.

Inter/National Coalition for Electronic Portfolio Research: <http://ncepr.org/>.



Irons, A. (2008). *Enhancing Learning Through Formative Assessment and Feedback*. Routledge.

JISC, HE Academy, and ALT (2008). Exploring Tangible Benefits of e-Learning. Retrieved in May 2010 from <http://www.jiscinfonet.ac.uk/publications/info/tangible-benefits-publication>.

JISC, (2009). Effective practice in a digital age. Retrieved in May 2010 from <http://www.jisc.ac.uk/publications/programmerelated/2009/effectivepracticedigitalage.aspx>.

JISC, (2007). Effective practice with e-assessment. Retrieved in May 2010 from [http://www.jisc.ac.uk/publications/programmerelated/2007/pub\\_eassesspracticeguide.aspx](http://www.jisc.ac.uk/publications/programmerelated/2007/pub_eassesspracticeguide.aspx).

JISC TechDis: Technology change for inclusion. 12 steps towards embedding inclusive practice with technology as a whole institution culture in UK Higher Education. Available from [www.techdis.ac.uk](http://www.techdis.ac.uk).

Jordan, L (2009). Transforming the student experience at a distance: designing for collaborative online learning. *Engineering Education* vol.4 issue 2 25-36.

Jordan, S. and Mitchell, T. (2009). e-assessment for learning? The potential of short-answer free-text questions with tailored feedback. *British Journal of Educational Technology*, 40(2), 371-385.

Juwah, C., Macfarlane-Dick, D., Matthew, B., Nicol, D., Ross, D., Smith, B. (2004). Enhancing student learning through effective formative feedback (SENLEF). The Higher Education Academy. Retrieved in June 2010 from [http://www.heacademy.ac.uk/resources/detail/resource\\_database/id353\\_effective\\_formative\\_feedback\\_juwah\\_etal](http://www.heacademy.ac.uk/resources/detail/resource_database/id353_effective_formative_feedback_juwah_etal).

Kirkpatrick, D. (1976). Evaluation of training. *Training and Development Handbook: A Guide to Human Resource Development*, 2, 18-11.

Kochakornjarupong, D. and Brna, P. (2010). Helping feedback-givers to improve their feedback. In D. Whitelock and P. Brna (eds) Special Issue 'Focusing on electronic feedback: feasible progress or just unfulfilled promises?' *Int. J. Continuing Engineering Education and Life-Long Learning*, Vol. 20, No. 2, pp.148-168.

Laurillard, D. (1993). *Rethinking University Teaching*, London: Routledge.

Lee, G. and Weerakoon, P. (2001). The role of computer-aided assessment in health professional education: a comparison of student performance in computer-based and paper-and-pen multiple-choice tests. *Medical Teacher*, Vol 23, No. 2, 152 - 157.

MacKenzie, Don (2003). Assessment for e-learning: what are the features of an ideal e-assessment system? 7th International CAA Conference, Loughborough. Retrieved in April 2010 from <http://www.caaconference.com/pastConferences/2003/proceedings/index.html>.

MacLellan, E. (2001). Assessment for Learning: The differing perceptions of tutors and students, *Assessment & Evaluation in Higher Education*, 26: 4, 307 — 318.

- Marriott, P. (2009). Students' evaluation of the use of online summative assessment on an undergraduate financial accounting module. *British Journal of Educational Technology*, 40(2), 237-254.
- McKenna, C. and Bull, J. (2000). Quality assurance of computer-assisted assessment: practical and strategic issues. *Quality Assurance in Education*. 8(1), 24-31.
- Middleton, A. and Nortcliffe, A. (2010). Audio feedback design: principles and emerging practice. In D. Whitelock and P. Brna (eds) Special Issue 'Focusing on electronic feedback: feasible progress or just unfulfilled promises?' *Int. J. Continuing Engineering Education and Life-Long Learning*, Vol. 20, No. 2, pp.208-223.
- Mitchell, T., Aldridge, N., Williamson, W. and Broomhead, P. (2003). Computer based testing of medical knowledge. 7th International CAA Conference, Loughborough. Last accessed September 2010 from <http://www.caaconference.com/pastconferences/2003/index.asp>.
- Nicol, D. (2007). Improving assessment after the task. Re-Engineering Assessment Practices in Scottish Education (REAP) Online Resources. Last accessed August 2010 at <http://www.reap.ac.uk/nss/nssAfter 02.html>.
- Nicol, D. (2007). Laying a foundation for lifelong learning: Case studies of e-assessment in large 1st-year classes. *British Journal of Educational Technology Vol 38 No 4* 668–678.
- Nicol, D. (2009). Assessment for learner self-regulation: enhancing achievement in the first year using learning technologies. *Assessment & Evaluation in Higher Education*, 34: 3, 335 — 352.
- Nicol, D. (2010). From monologue to dialogue: improving written feedback processes in mass higher education. *Assessment & Evaluation in Higher Education*, 35: 5, 501 — 517.
- Nicol, D. J. and Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: a model and seven principles of good feedback practice. *Studies in Higher Education*, 31(2), 199-218.
- Nyquist, J. B. (2003). The benefits of reconstruing feedback as a larger system of formative assessment: A meta-analysis. Unpublished master of science thesis, Vanderbilt University.
- PASS-IT (2004). Good Practice Guide in Question and Test Design, (extended from "Designing and Using Objective Tests" by Dr Joanna Bull & Dr Colleen McKenna). <http://www.pass-it.org.uk/resources/031112-goodpracticeguide-hw.pdf>.
- Ricketts, C. and Wilks, S. J. (2002). Improving student performance through computer-based assessment: insights from recent research. *Assessment & Evaluation in Higher Education*, 27(5) 475.
- Sandene, B., Horkay, N., Bennett, R. E., Allen, N., Braswell, J., Kaplan, B. and Oranje, A. (2005). Online Assessment in Mathematics and Writing: Reports from the NAEP Technology-Based Assessment Project, Research and Development Series. Last accessed August 2010 from <http://nces.ed.gov/nationalreportcard/pubs/studies/2005457.asp>.
- Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, 18 (2), 119-144.

- Sadler, D. R. (1998). Formative assessment: revisiting the territory. *Assessment in Education*, 5(1) 77-84.
- Shephard, K., Warburton, B., Maier, P. and Warren, A. (2006). Development and evaluation of computer-assisted assessment in higher education in relation to BS7988. *Assessment & Evaluation in Higher Education*, 31: 5, 583 — 595.
- Sieber, V. (2009). Diagnostic online assessment of basic IT skills in 1<sup>st</sup>-year undergraduates in the Medical Sciences Division, University of Oxford. *British Journal of Educational Technology*, 40 (2), 215 -226.
- Sim, G., Holifield, P. and Brown, M. (2004). Implementation of computer assisted assessment: lessons from the literature, *ALT-J*, 12: 3, 215 — 229.
- Sloan, D. and Walker, D. (2008). Evaluating the accessibility to disabled people of e-Assessment tools. 12th CAA International Computer Assisted Assessment Conference (2008) (Eds. F. Khandia), Professional Development. 8 and 9 July, Loughborough University. Also available from <http://www.caaconference.com/pastConferences/2008/proceedings/index.html>.
- Sotto, E. (2007). *When Teaching Becomes Learning: A Theory and Practice of Teaching*. Second Edition. London, Continuum International Publishing Group.
- Stephens, Derek, Bull, Joanna and Wade, Winnie (1998). Computer-assisted Assessment: suggested guidelines for an institutional strategy. *Assessment & Evaluation in Higher Education*, 23: 3, 283 — 294.
- Strang, K. D. (2010). Measuring self regulated e-feedback, study approach and academic outcome of multicultural university students. In D.Whitelock and P.Brna (eds) Special Issue 'Focusing on electronic feedback: feasible progress or just unfulfilled promises?' *Int. J. Continuing Engineering Education and Life-Long Learning*, Vol. 20, No. 2, pp.239-255.
- Strivens, J (2006). Efficient Assessment of Portfolios. Retrieved in May 2010 from <http://cetl.open.ac.uk/pbpl/pics/d58920.pdf>.
- Thomas, R., Ashton, H., Austin, B., Beevers, C., Edwards, D. and Milligan, C. (2004). Assessing Higher Order Skills Using Simulations. Proceedings of the 8th International CAA Conference, Loughborough University Retrieved in May 2010 from [http://www.caaconference.com/pastConferences/2004/proceedings/Thomas\\_R.pdf](http://www.caaconference.com/pastConferences/2004/proceedings/Thomas_R.pdf).
- Walker, D., Adamson, M. and Parsons, R. (2004). Staff education, learning about online assessment, online. Proceedings of the 8th International Computer Assisted Assessment Conference. Loughborough. Retrieved in June 2010 from [http://www.caaconference.com/pastConferences/2004/proceedings/Walker\\_Parsons.pdf](http://www.caaconference.com/pastConferences/2004/proceedings/Walker_Parsons.pdf).
- Warburton, B. (2009). Quick win or slow burn: modelling UK HE CAA uptake, *Assessment & Evaluation in Higher Education*, 34: 3, 257 — 272.
- Whitelock, D. (2006a). e-Assessment: Case Studies of Effective and Innovative Practice. Final report for JISC ITT funded project conducted by The Open University and University of Derby. Retrieved in April 2010 from <http://kn.open.ac.uk/public/document.cfm?docid=10817>.

Whitelock, D. (2006b). Computer Assisted Formative Assessment (CAFA): An Investigation into the Pedagogical Push and Technological Pull for Designing Holistic e-Assessment. Final report of CAFA Project. Retrieved in April 2010 from <http://kn.open.ac.uk/public/document.cfm?docid=10798>.

Whitelock, D. (2006). Electronic Assessment: Marking, Monitoring and Mediating Learning. In McAndrew, P. and Jones, A. (eds) *Interactions, Objects and Outcomes in learning*. Special Issue of *International Journal of Learning Technology*. Vol. 2 Nos. 2/3 p.264-276.

Whitelock, D. (2009). Editorial: e-assessment: developing new dialogues for the digital age. *British Journal of Educational Technology*, 40(2), 199-202.

Whitelock, D. (2010). Activating Assessment for Learning: are we on the way with Web 2.0? In M.J.W. Lee & C. McLoughlin (Eds.) *Web 2.0-Based-E-Learning: Applying Social Informatics for Tertiary Teaching*. IGI Global.

Whitelock, D. and Brasher, A. (2006). Developing a Roadmap for e-assessment: which way now? Proceedings of 10th CAA International Computer Assisted Assessment Conference, Loughborough University, 4/5 July 2006 edited by Myles Danson ISBN 0-9539572-5-X pp.487-501. Also available from <http://www.caaconference.com/pastConferences/2006/proceedings/index.html>.

Whitelock, D., Mackenzie, D., Whitehouse, C., Ruedel, C. and Rae, S. (2006). Identifying innovative and effective practice in e-assessment: findings from 17 UK case studies. 10th CAA International Computer Assisted Assessment Conference, Loughborough University, 4/5 July 2006 edited by Myles Danson ISBN 0-9539572-5-X pp.505-511. Also available from <http://www.caaconference.com/pastConferences/2006/proceedings/index.html>.

Whitelock, D., Ruedel, C. and Mackenzie, D. (2006). e-Assessment: Case Studies of Effective and Innovative Practice. Final report for JISC ITT funded project conducted by The Open University and University of Derby. Retrieved in May 2010 from <http://kn.open.ac.uk/public/document.cfm?docid=10817>.

Whitelock, D., and Watt, S. (2007). Supporting Tutors with their feedback to students. In F. Khandia (Ed.), 11th International Computer Assisted Assessment Conference (pp. 421-432). Loughborough: Loughborough University. Also available from <http://www.caaconference.com/pastConferences/2007/proceedings/index.html>.

Whitelock, D. and Watt, S. (2008). Reframing e-assessment: adopting new media and adapting old frameworks. *Learning, Media and Technology*, Vol. 33, No. 3, September 2008, pp.153–156 Routledge, Taylor & Francis Group. ISSN 1743-9884.

Whitelock, D., Watt, S., Raw, Y. and Moreale, E. (2003). Analysing Tutor Feedback to Students: First steps towards constructing an Electronic Monitoring System. *Association for Learning Technology Journal (ALT-J)*. Vol. 11, No. 3, pp. 31-42 ISSN 0968-7769.

William, D. (2008). Balancing Dilemmas Traditional Theories and New Applications. In A. Haines and L. McDowell (Eds) *Balancing Assessment and Learning in Contemporary Education* pp.267-281. Taylor and Francis Group, Routledge.

William, D. (2009). Assessment for learning: why, what and how? An inaugural professorial lecture by Dylan William. Institute of Education, University of London.

Williams, J. B. and Wong, A. (2009). The efficacy of final examination: A comparative study of closed-book, invigilated exams and open-book, open-web exams. *British Journal of Educational Technology Vol 40 No 2* 227–236.

Wills, G., Gilbert, L., Warburton, B. and Gale, V. (2010). Writing effective e-assessments: Hints and tips from practitioners. Output of the JISC EASiHE project available from <http://easihe.ecs.soton.ac.uk/>.

Winne, P. H., Nesbit, J. C., Kumar, V., Hadwin, A. F., Lajoie, S. P., Azavedo, R. and Perry, N. E. (2006). Supporting self-regulated learning with gStudy software: the learning kit project. *Technology, Instruction, Cognition and Learning, Vol. 3, Nos.1/2*, pp.105-113.

Zakrzewski, S. and Steven, C. (2000). A Model for Computer-based Assessment: The catherine wheel principle. *Assessment & Evaluation in Higher Education, 25: 2*, 201 — 215.

Zakrzewski, S. and Steven, C. (2003). Computer-based assessment: quality assurance issues, the hub of the wheel. *Assessment & Evaluation in Higher Education, 28: 6*, 609 — 623.