eFEP Final Evaluation Report

How to cite:
Fernez-Toro, Maria; Furnborough, Concha and Polisca, Elena (2013). eFEP Final Evaluation Report. JISC.

For guidance on citations see FAQs.

Version: Version of Record

Link(s) to article on publisher’s website:
http://repository.jisc.ac.uk/id/eprint/5389

Copyright and Moral Rights for the articles on this site are retained by the individual authors and/or other copyright owners. For more information on Open Research Online’s data policy on reuse of materials please consult the policies page.
Executive summary

1.1 Brief description of the eFEP project

The e-Feedback Evaluation Project (eFEP) evaluated the use of spoken and written e-feedback in a context in which these modes of delivery had been adopted by a Higher Education institution (the Open University) across an entire subject area, in this instance Modern Foreign Languages. The project undertook a systematic analysis of the content of the feedback itself, how feedback content varied from one tutor to another depending on the medium used, and whether students engaged in different ways with specific features of the feedback they received.
1.2 Evaluation questions and purpose of evaluation

Our evaluation questions focused on three areas:

1. **Staff and student attitudes and perceptions:**
   How do staff and students perceive different forms of e-feedback, specifically feedback summary forms, tutor’s annotations on student scripts, and audio-recorded feedback?

2. **Quality of the feedback given by tutors:**
   What aspects do tutors cover and how rich is their feedback in each of these e-feedback media?

3. **Student engagement with the feedback:**
   How do students engage with the feedback given by tutors on their assignments?

A smaller-scale evaluation was also conducted at the University of Manchester (UoM) to evaluate the extent to which the lessons learnt at the Open University (OU) are applicable to non-distance learning contexts.

1.3 Brief summary of evaluation methodology

A range of data collection methods were used in order to evaluate these three key areas:

The **quality of the feedback given by tutors** was evaluated using the Feedback Analysis Chart for Tutors (FACT), an evaluation tool developed for this project, initially based on the notion of feedback ‘depth’ in Brown and Glover (2006) and subsequently expanded. All tutors’ comments and annotations were coded in terms of their orientation (whether they focused on strengths and/or weaknesses), the performance areas on which they focused (e.g. accuracy and range of the language used, structure, accuracy and relevance of the content), and the depth of the feedback given (i.e. whether it provided indications, corrections, explanations, or advice for future assignments). The sample comprised a total of 216 assignments, half of which were written assignments (scripts with tutors’ annotations and online feedback summary forms) and the other half audio–recorded feedback on speaking assignments, also with online feedback summary forms.

**Staff and students’ attitudes and perceptions** were evaluated by means of online questionnaires completed by 96 tutors and 736 students, while **student engagement with the feedback** was evaluated through a ‘feedback on feedback’ exercise where 10 OU students recorded screencasts as they talked through the written feedback they had received on one of their writing assignments. Audio-recorded peer-to-peer feedback was also collected from 58 UoM students.

1.4 Summary of main findings, and implications of these

In terms of **attitudes and perceptions** OU students (and tutors) generally found written feedback more useful than audio feedback, which was used primarily for pronunciation and intonation feedback on speaking assignments. At UoM, where
audio-feedback is not routinely used, students responded very positively to a more personalised approach to receiving feedback. This indicated how audio-feedback can generate new possibilities in terms of student engagement with feedback, active dialogue between tutors and students, and internalisation of feedback by students.

Statistical comparison of the nature and quality of tutor use of feedback revealed that the feedback summary forms (for both written and spoken assignments) contained proportionately more comments on strengths than on weaknesses, as well as on the content and structure of the assignment rather than use of language. By contrast, in written annotations and audio feedback a higher proportion of comments related to use of language rather than content, and to weaknesses than to strengths. Overall, feedback on language use is deeper than on content, and feedback related to weaknesses is also deeper than feedback related to strengths.

The FACT method generated clearly quantified profiles of tutors’ feedback. However, this was more complex than anticipated, and so from a training perspective it served as a stimulus for reflection rather than visual evidence of individual profiles, as originally envisaged. The pilot feedback alignment exercise did not stimulate tutors to try it out themselves and we are discussing the development of a prototype online delivery method as a possible follow-up.

Data on student engagement with the feedback indicated cognitive responses to feedback on weaknesses leading to active integration or informed acceptance of tutor feedback. Positive affective responses to feedback on strengths occurred when students perceived an acknowledgement of their efforts or the establishment of a personal rapport with them. Metacognitive responses to feedback on strengths led to planning strategies for improving future performance. However, unexplained corrections to more complex syntactical errors led to ineffective feedback dialogue, and at more advanced levels unexplained lexical corrections were sometimes rejected on both cognitive and affective grounds by students perceiving the tutor as failing to recognise their creative experimentation with the language.

The findings confirm tutors’ fears that their feedback does not always achieve its intended purpose, for reasons which the project identified as either cognitive (e.g. due to insufficient depth of feedback) or affective (apparent lack of acknowledgment of students’ efforts or individuality). Since the participants in this study were highly motivated achievers this is likely to occur more frequently across the spectrum as a whole.

Peer-to-peer feedback, as piloted in UoM, also demonstrated students’ ability to reproduce feedback-specific vocabulary in an appropriate and context-specific manner, indicating engagement with the feedback received from their tutors. Follow-up of peer-to-peer evaluation is planned.

In the light of these findings, the following aspects should be considered in OU guidelines and training for language tutors:

- broadening the scope of audio feedback beyond pronunciation alone
Assessment and Feedback programme:  
The e-Feedback Evaluation Project (eFEP)  
Final Evaluation Report

- using FACT-based criteria as a checklist for reflection on their own use of feedback
- giving deeper feedback (e.g. examples, explanations) on strengths as well as on weaknesses
- promoting student engagement and feedback dialogue (e.g. feedback alignment exercise, feedback on feedback)

1.5 Summary of recommendations

- Overall, tutors saw e-Feedback as an improvement over conventional feedback, in terms of the quality both of the students’ learning experience and of the tutors’ teaching experience. For obvious logistical reasons returning to paper-based feedback at the OU is not an option, so it is reassuring to see that e-feedback also has benefits in terms of pedagogical value.

- OU language tutors are routinely instructed to use audio mostly for feedback related to pronunciation and intonation feedback in speaking assignments. However at UoM, where audio-feedback was not routinely used, students responded very positively to the more personalised approach that audio feedback offered. Feedback analysis also showed that, although it dealt with fewer points, audio feedback from OU tutors contained significantly more explanations than their feedback in written media. This suggests that the OU may not be making the most of the medium’s affordances and tutors could be encouraged to focus on a wider range of skills rather than restricting audio feedback to matters of pronunciation.

- The FACT analysis revealed a consistent tendency to give deeper feedback on weaknesses than on strengths. This may well respond to a genuine need and therefore one has to be cautious before making specific recommendations, however the pattern must be noted and critically evaluated by tutors and line managers with specific responsibility for the quality of teaching through feedback at the OU. A working group comprising members of the eFEP team, staff tutors and central academic staff, has been put in place in order to identify which of the observed patterns are justified, which ones may need to change, and how the briefing and training of OU tutors and monitors can be improved in the light of this evaluation.

- Project findings clearly demonstrate the potential of feedback on feedback and peer-to-peer feedback for promoting learner engagement with assignment feedback, and we recommend that they are both integrated at some point into the curriculum. Tutors may also wish to use them at their discretion, possibly at the beginning of a module for diagnostic purposes or when they feel that a student does not appear to learn from the feedback provided.
2. Background and context

2.1 Purpose of the evaluation and core evaluation questions

The aim of the project was to evaluate the use of spoken and written e-feedback in a context in which these modes of delivery had been adopted by a Higher Education institution across an entire subject area. One such context was the Open University, where the use of both audio-recorded and written e-feedback had been standard practice at the Department of Languages for a number of years. The evaluation looked at staff and student perceptions of assignment feedback, the quality of feedback itself, and student engagement with the feedback.

More specifically, the project aimed to evaluate:

- the students’ and tutors’ attitudes to/perceptions of assignment feedback in each of the media commonly used at the OU
- the quality of feedback in three of the media used in terms of the criteria being assessed and the depth of feedback on strengths and weaknesses.
- the effectiveness of feedback in terms of student engagement and response

The purpose of this evaluation was to find out students’ and tutors’ perceptions on assignment feedback, to understand what use they make of it, and identify ways of improving the quality of the feedback provided, the effectiveness of the different media used, and the efficiency of the feedback process.

In order to evaluate the extent to which the lessons learnt at the OU are applicable to non-distance learning contexts a smaller-scale evaluation was also conducted at the University of Manchester.

The core evaluation questions in each of these three areas were:

1. **Staff and student attitudes and perceptions:**
   What do staff and students think about assignment feedback and how useful do they find different forms of e-feedback such as feedback summary forms, tutor’s annotations on student scripts, and audio-recorded feedback?

2. **Quality of the feedback given by tutors:**
   What aspects do tutors cover and how rich is their feedback in each of these e-feedback media?

3. **Student engagement with the feedback:**
   How do students engage with the feedback given by tutors on their assignments and how could feedback be improved in order to promote and enhance learner engagement?
2.2 Description of the eFEP Project and its context

Over recent years, JISC has funded a number of projects that evaluate different uses of technology as a means of providing richer, more effective feedback without increasing staff workloads. Whilst research findings to date are encouraging, notably in the area of audio-recorded e-feedback, the methods evaluated have tended to be at an experimental stage, focusing on relatively small samples and individual modules (see section 2.4 below). As a result there is a pressing need for evaluative studies that are set in contexts in which:

- different types of e-feedback (supported by written and spoken media) are used systematically and consistently across an entire programme;
- giving and receiving e-feedback is a familiar experience for staff and students alike;
- well-established standard procedures are in place for training and supporting tutors in using e-feedback, and for monitoring the quality of the feedback that students receive.

These conditions are met by the Open University Languages programme and provide enough critical mass to allow for a more systematic approach to evaluating e-feedback.

Another issue relates to the aspects of e-feedback upon which most evaluative studies have been concentrated so far. Whilst a number of important criteria such as staff and student preferences and perceptions, and practical and technical issues, have already been well covered in previous evaluative studies, a more systematic analysis was needed with regard to the content of the feedback itself, how feedback content varied from one tutor to another depending on the medium used, and whether students engaged in different ways with specific features of the feedback they received. This project therefore used systematic feedback analysis and observational methods in addition to self-reported data from students and tutors.

The key objectives of the project were therefore to:

- examine the ways in which students and tutors actually use spoken and written e-feedback when these have become standard practice across an entire programme;
- evaluate the perceptions and preferences of tutors and students in relation to spoken and written e-feedback in those learning contexts;
- investigate the ways in which students engage with the written and spoken e-feedback that they receive;
- evaluate the extent to which the findings from the above evaluations apply to non-distance learning institutions in which e-feedback is an emerging practice;
- identify the areas in which further support is needed in order to maximise the benefits of spoken and written e-feedback;
- produce a set of training and support tools for students and tutors.
2.3 Target population for the eFEP and relevant stakeholders for the evaluation

Both internal and external stakeholders were identified at the start of the project (see Tables 1a and 1b).

Table 1a. Internal stakeholders and their interests.

<table>
<thead>
<tr>
<th>Internal stakeholders</th>
<th>Stake / Interest</th>
</tr>
</thead>
<tbody>
<tr>
<td>The Open University</td>
<td>Maintaining the lead in innovative pedagogy, e-learning and student support in open learning environments, notably in the area of academic feedback.</td>
</tr>
<tr>
<td>Faculty</td>
<td>Developing more time- and resource-efficient assessment processes. An Assessment Working Group was recently created in DoL to address these issues in the light of current budget cuts.</td>
</tr>
<tr>
<td>Module Developers at DoL</td>
<td>Developing assessment strategies that promote learning dialogue, encourage students to engage with tutor feedback, and make the most out of the technology available</td>
</tr>
<tr>
<td>Teaching staff</td>
<td>Providing their students with feedback that makes the most out of the media available, without putting unreasonable demands on their workload. Access to training resources that help them capitalise on the strengths of each medium whilst ensuring that all the functions of effective feedback are duly covered</td>
</tr>
<tr>
<td>Staff Tutors (tutors’ line managers)</td>
<td>Using the project findings to inform the induction and training of new teaching staff, and staff development activities for which staff tutors are responsible. The new issues raised will encourage discussion and cross-fertilisation among fellow practitioners.</td>
</tr>
<tr>
<td>Monitors</td>
<td>Evaluating the feedback provided by language tutors and providing advice that is informed by the findings of this project.</td>
</tr>
<tr>
<td>Students</td>
<td>Accessing feedback that meets their needs and makes them want to engage with it in order to learn. Online training events based on the findings should also contribute to develop a sense of learner’s ownership in relation to feedback. Taking part in observational studies (as recipients or providers of e-feedback), should be an empowering experience for the students involved.</td>
</tr>
<tr>
<td>The University of Manchester</td>
<td>To promote the adoption of e-learning practices beyond the modules currently covered, with the aim of enhancing the student experience. To remain at the forefront of e-learning and didactic practices nationally.</td>
</tr>
</tbody>
</table>

Table 1b. External stakeholders and their interests.

<table>
<thead>
<tr>
<th>External stakeholders</th>
<th>Stake / interest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teaching institutions in the UK and beyond</td>
<td>Accessing the lessons learnt by a large institution in which the use of written and spoken media to provide tutor e-feedback has been fully integrated in the curriculum and assessment processes for a number of years; improving their student satisfaction scores related to feedback on the NSS.</td>
</tr>
<tr>
<td>JISC Programme</td>
<td>Supporting the project to achieve the project outputs; encouraging cross-fertilisation between project teams.</td>
</tr>
<tr>
<td>JISC Community</td>
<td>Accessing the outputs of the project and engaging with the issues raised.</td>
</tr>
</tbody>
</table>

2.4 Related work/studies

In recent years, feedback has been at the forefront of pedagogical research as a means to enhance the student experience whilst exploring the affordances of, and keeping up with, technological innovations. Numerous studies have thus focussed on various
aspects of e-feedback, notably audio/visual feedback (AFAL, ASSET, Sounds good). Other studies have explored ways in which feedback may become more meaningful for students by integrating, for example, peer-to-peer practices (PEER) or blogs (MAC) within a given module. Other studies and projects have looked at ways in which feedback would become more interactive (InterACT, eAssessment in HE, ESCAPE) and streamlined for users (Institutional submissions & management system for assessment of open assignment). e-FEP has capitalised on the experience of such studies in terms of added value: whilst these studies have mostly focussed on one aspect of the feedback process and investigated the outcomes of one innovation within a given curriculum, e-FEP has moved the focus on to evaluating different aspects of e-feedback when innovations such as the ones mentioned above have become standard practice within one institution. In this respect, e-FEP has been able to acknowledge, and learn from, the experiences of previous studies whilst adding an evaluative dimension to the current literature.

Below is a selected list of studies which have affinities with the e-FEP project and have informed its methodology.

- AFAL: Optimising Audio Feedback Assisted Learning for Student and Staff Experience [http://www.jisc.ac.uk/whatwedo/programmes/elearning/ltig/afal.aspx]
- ASSET: moving forward through feedback [http://www.reading.ac.uk/asset/]
- eAssessment in Higher Education [http://www.jisc.ac.uk/whatwedo/programmes/institutionalinnovation/easihe.aspx]
- Effecting sustainable change in assessment practice and experience (ESCAPE) [http://www.jisc.ac.uk/whatwedo/programmes/elearning/curriculumdelivery/escape.aspx]
- InterACT - Interactive Assessment and Collaboration via Technology [http://www.jisc.ac.uk/whatwedo/programmes/elearning/assessmentandfeedback/interACT.aspx]
- Making Assessment Count (MAC) [https://sites.google.com/a/staff.westminster.ac.uk/mac/]
- Sounds good: quicker, better assessment using audio-feedback [http://sites.google.com/site/soundsgooduk/]

3 Evaluation approach
The project evaluated three core areas related to e-feedback on language assignments: staff and student attitudes and perceptions, the quality of the feedback itself, and student engagement with the feedback.

In order triangulate the findings, the methods used comprised a combination of qualitative and quantitative approaches using self-reported as well as observational data.
3.1 Design of the evaluation

3.1.1 Staff and student attitudes and perceptions
Conventional evaluation methods based on questionnaires and self-reported data were used in order to investigate the perceptions, attitudes and preferences of students and tutors. Two separate surveys were produced at each institution for tutors and students. In order to facilitate comparisons between the two groups, both questionnaires followed a similar structure. Baseline data was collected from those tutors who had experience of using paper and cassette-based feedback at the Open University before electronic forms of delivery were introduced in the Department of Languages. These tutors were asked to rate the two forms of delivery in terms of the quality of the student’s learning experience and the quality of the tutor’s teaching experience.

3.1.2 Quality of the feedback itself
In order to evaluate the quality of the feedback itself, a specific analysis framework was developed, based initially on Brown and Glover’s (1986) notion of feedback ‘depth’ and subsequently refined and expanded to include five levels of scaffolding that tutors may provide in their feedback on an assignment’s strengths or on its weaknesses. This formed the basis of the Feedback Analysis Chart for Tutors (FACT) evaluation tool which was developed for this project. A screencast description of the tool is available on the project website. The FACT tool was designed to provide observational evidence that would tell us more specifically:

- what features characterise the actual feedback that tutors give through each of the spoken and written media at their disposal;
- how students respond to, and engage with, the particular features that are present in the actual feedback that they receive.

3.1.3 Student engagement with the feedback
Two different approaches were used in order to evaluate student engagement:

The first approach consisted of asking students to talk us through the feedback they got on one of their assignments and examine their responses. This method, which could be referred to as feedback on feedback, was used with 10 students enrolled on Spanish modules at the Open University. The students’ self-reported data were examined alongside observational data drawn from a FACT analysis of the feedback that they were commenting upon.

The second approach consisted of analysing the audio-recorded feedback that students gave to other students at different stages in their degree. This peer to peer feedback method was adopted at the University of Manchester, where it is standard practice within Italian modules (see peer to peer feedback questionnaire). The scope of the project only allowed for a qualitative analysis of the peer-to-peer feedback within the duration of the project. However the students’ feedback was encoded in the same way as feedback from tutors using the FACT tool and is therefore available for further quantitative analysis.
3.2 **Data collection and analysis**

### 3.2.1 Tutor and student surveys

**Open University:**

Online surveys were designed by the project team and then converted for online delivery by the Open University’s Institute of Educational Technology (IET), who also advised on sample selection and questionnaire design.

- **Student survey**

  Table 2. Student survey participants according to language and level.

<table>
<thead>
<tr>
<th>Language</th>
<th>Beginner</th>
<th>Lower Int.</th>
<th>Upper Int.</th>
<th>Advanced</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chinese</td>
<td>27</td>
<td></td>
<td></td>
<td></td>
<td>27</td>
</tr>
<tr>
<td>French</td>
<td>106</td>
<td>72</td>
<td>23</td>
<td>31</td>
<td>232</td>
</tr>
<tr>
<td>German</td>
<td>63</td>
<td>41</td>
<td>16</td>
<td>9</td>
<td>129</td>
</tr>
<tr>
<td>Italian</td>
<td>85</td>
<td>46</td>
<td></td>
<td></td>
<td>131</td>
</tr>
<tr>
<td>Spanish</td>
<td>90</td>
<td>80</td>
<td>18</td>
<td>22</td>
<td>210</td>
</tr>
<tr>
<td>Welsh</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td>7</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>378</strong></td>
<td><strong>239</strong></td>
<td><strong>57</strong></td>
<td><strong>62</strong></td>
<td><strong>736</strong></td>
</tr>
</tbody>
</table>

- **Tutor survey: N=96 (same modules as above)**

Given the small size of the project team, a full, systematic comparison between the results of both surveys was not possible within the time available for this phase of the project. Therefore data analysis focused on comparing tutors’ and students’ responses to a selection of key questions from the survey. These related to:

- What students regard as the most important aspects in assignment feedback
- Which of the three media students and tutors find most / least useful
- Responses to 15 statements about feedback in general (using Likert scales)

The rest of the dataset remains available for future analysis.

**University of Manchester:**

A similar student survey involving students enrolled on Italian modules was conducted in Manchester using paper questionnaires. The sample comprised 58 students from four different levels: 17 beginners, 16 post-beginners, 9 post A-level, 16 finalists.

### 3.2.2 Analysis of tutor feedback

A sample of assignment feedback from Spanish modules at all four levels taught at the Open University was collected, covering beginner, lower intermediate, upper intermediate, and advanced levels. The sample included 9 tutors from each level, and 3 students per tutor. For each of these students the sample collected included the full
set of e-feedback provided on one oral assignment and one written assignment. The sample comprised a total of 216 assignments, consisting of the following:

For written assignments:
- 108 online feedback summary forms (internally known as PT3 forms)
- 108 written assignment scripts with tutors’ annotations on Word markup.

For speaking assignments:
- 108 online feedback summary forms (PT3 forms)
- 105 audio–recorded feedback on speaking assignments (MP3 files returned by tutors)

The dataset was anonymised and the audio files transcribed. A sample of 6 sets was analysed independently by a team member and an appointed consultant and subsequently tested for inter-rater reliability. As it was the first time the FACT tool had been used a series of adjustments were required. After three rounds of tests a highly reliable version was produced, with an inter-rater reliability index of 0.97. Full notes on coding conventions with examples based on language assignments are available on the project website.

3.2.3 Feedback on feedback
Out of the 736 language students who took the student survey, the 210 who were taking Spanish modules were invited to participate in this follow-up study. 88 agreed to be contacted and were sent an overall description of the study. Twenty of them subsequently requested the necessary instructions to produce the recordings. The final sample consisted of 10 students, who were the only ones to return a set of recordings. Such a high dropout rate was expected given the voluntary nature of the task, the challenge of trying out a new technological tool, and the fact that these were adult distance learners who had just completed their respective modules. As a result, the sample is not entirely representative of the student population as a whole, but of a highly motivated, high-achieving minority. Indeed their marks on the assignment used all ranged between 75% and 94%, and this was taken into consideration when interpreting the data. All the levels taught at the OU were represented in the sample, which consisted of two students taking the beginner module, two from the lower intermediate module, four from the upper intermediate, and two from the advanced module. The sample comprised 5 males and 5 females.

Students were given a written set of instructions and a screencast showing a simulated talk-through recorded by one of the researchers. All the necessary material was available online. The recording tool used was Jing, which allows a maximum recording time of 5 minutes. Students were asked to produce two recordings each: one about their marked written script, known as a Tutor Marked Assignment (TMA), and another one about the accompanying feedback summary form (PT3). Students were sent anonymised copies of these document files so that no personal details could be seen on their recordings. In their task brief, they were encouraged to talk us through the assignment feedback, covering any aspects that they considered relevant, such as their first reaction to the feedback, which comments they did or did not understand, which ones they found useful or not useful, what feelings different comments elicited, what use students made of the feedback, and what they learned from it. Once the
recordings were completed, students submitted them by email. Thus, from the initial briefing to the final submission, the entire process took place electronically.

Each student’s recordings were analysed in terms of their use of the two media (TMA script and PT3 form); their cognitive, affective and metacognitive responses to comments on strengths and comments on weaknesses; and their responses to different depths of feedback relating to strengths and weaknesses of their work.

3.2.4  Peer-to-peer feedback
Peer-to-peer (p2p) feedback was collected from two Italian modules at the University of Manchester:
•  beginners module, where p2p feedback was a compulsory component of the last piece of assessed coursework in semester 2.
•  advanced module, where p2p was a recommended activity conducted on weeks 4 and 9 of semester 2.

The sample collected comprised one audio recording in which one group of students from the beginner module gave spoken feedback on a composition written by another group, and 22 recordings of advanced students giving spoken feedback on individual compositions written by their peers.

Spoken feedback from the students’ tutors was also collected for comparison purposes. A preliminary qualitative analysis was conducted comparing tutors’ and students’ approaches to each of four criteria (morphology and syntax, lexis and register and content), plus any additional comments. Feedback from 17 students to their peers has also been encoded using the FACT tool and is available for further quantitative analysis in future.

3.3  Evaluation limitations (e.g. related to methods, data sources, biases)

3.3.1  Tutor and student surveys
The two questions ‘Which of these media do you find most useful?’ and ‘Which of these media do you find least useful?’ were too general and did not allow responses whereby both were equally useful.

Only a part of the data collected could be analysed within the life of the project, however the team will continue the work in preparation for a research article to be submitted for publication in 2014.

3.3.2  The FACT tool
While FACT is a useful research tool in terms of identifying overall patterns of use of different feedback media, coding requires complex guidelines in order to be reliable. This implies that it is not suitable for quantitative evaluation by practitioners. For training purposes, we adopted a coarser, much simpler approach to FACT coding, as clarity is more important than high inter-rater reliability when the tool is used for self-evaluation.
It must also be stressed that the results of any FACT analysis must be interpreted with caution. For example, ‘deeper’ feedback is not necessarily the most appropriate in all contexts, and over-prescriptive approaches aiming at some ideal FACT profile should be avoided.

3.3.3 Feedback on feedback
This aspect of the evaluation has two limitations: Firstly, the self-selected nature of the sample means that it did not represent the student population as a whole, and the study would need to be repeated with a randomly selected sample including less motivated and able students. Secondly, as the ‘feedback on feedback’ exercise conducted here was intended for research purposes, the students were addressing the researchers rather than their tutors, thus missing out on a valuable opportunity for genuine feedback dialogue. Despite these limitations, the fact that recordings were submitted at all shows that the method is potentially viable and could be implemented as a means of promoting feedback dialogue between students and tutors, both in face-to-face and distance learning environments.

3.3.4 Peer-to-peer feedback
Due to illness, this aspect of the evaluation could not be fully exploited. However preliminary results are promising and participants in various dissemination events have shown much interest in this approach. Therefore we are compiling a dataset in which FACT coding will be transferred to NVivo. Such a dataset has considerable potential for further qualitative analysis, including a comparison between the feedback given by 6 tutors and their students.

4 Evaluation findings

4.1 Staff and student attitudes and perceptions
How do staff and students perceive different forms of e-feedback, notably feedback summary forms, tutor’s annotations on student’s scripts, and audio-recorded feedback?

A comparison of survey results among Manchester and OU students shows that students at Manchester tend to find audio feedback more useful than written feedback, whereas OU students (and tutors) generally find written feedback more useful than audio feedback (see Figures 1-6).

Figure 1. Most useful feedback medium: OU tutor survey.
Figure 2. Most useful feedback medium: OU student survey.

Figure 3. Medium feedback preferences: Beginners UoM students.

Note: In Figures 3 to 6 SLLC refers to the feedback form used by the School of Languages, Linguistics and Cultures at the UoM.
Figure 4. Medium feedback preferences: Yr 2 Post-Beginners UoM students.

Figure 5. Medium feedback preferences Yr 2 Post-A level UoM students.

Figure 6. Medium feedback preferences: Finalists UoM students.
Audio-feedback has been used by the OU for a number of years to give feedback on oral tasks to its students. In particular, this feedback related to aspects of the spoken language, such as intonation and pronunciation. The adoption of audio-feedback by the OU proved to be the most obvious choice to provide feedback on speaking assignments to its distance-learning students, notably on pronunciation and intonation where written feedback is not an option: audio-feedback represented, in this instance, a practical solution to a logistical problem. On the other hand, Italian Studies at UoM does not use audio-feedback routinely. This mode of feedback delivery was chosen in order to enhance the student experience and at the same time to make better use of the time the tutor spent writing feedback forms. In this case, audio-feedback provided a more personalised approach to receiving feedback and students responded very positively to this innovation. At UoM language staff are encouraged to compile feedback forms with each piece of assessed coursework, print them and staple them to the coursework returned to the students. Overall, this is a very time-consuming approach, and the Manchester member of the team piloted audio-feedback with the dual purpose of trying to engage students in the feedback process by offering more personalised feedback whilst trying to optimise her use of time (cf. Sounds Good: Quicker, better assessment using audio-feedback project for recommendations on how to save assessors’ time).

Another possible reason is that the written feedback given to OU students is already extremely comprehensive in comparison with the written feedback given to UoM students. This is because UoM students also have access to face-to-face feedback in class. Indeed most of beginner and post-A level students in Manchester rated face-to-face as the most useful medium. Nevertheless, most post-beginners and finalists preferred audio, perhaps because in face-to-face learning contexts it elicits a more personalised and detailed feedback than feedback given in class.

Despite the fact that the two institutions are using audio-feedback for different purposes, it became clear through the project that audio-feedback can indeed open up a whole new world of possibilities in terms of: a) student engagement with the feedback process; b) active dialogue between two of the major stakeholders (tutors and students), and c) internalisation of feedback on the part of the students.

Tutors’ evaluation of eFeedback in comparison to conventional feedback

Tutors who had experienced paper-based as well as electronic feedback whilst being employed at the OU responded as follows:

- In terms of the quality of the students’ learning experience, 45 tutors stated that eFeedback was better than conventional feedback, 35 stated that both were just as good, and only 4 stated that eFeedback was not as good as conventional feedback. The advantages mentioned included clarity, speed and convenience, while the disadvantages related to technical problems which might cause anxiety.
- In terms of the quality of their own teaching experience, 47 tutors stated that eFeedback was better than conventional feedback, 28 stated that both were just as
Assessment and Feedback programme:  
The e-Feedback Evaluation Project (eFEP)  
Final Evaluation Report

good, and 9 stated that eFeedback was not as good as conventional feedback. The advantages mentioned included greater versatility, offering scope for using different media, while the disadvantages related to time required, and preference for not working on screen.

4.2 The quality of the feedback itself

What aspects do tutors cover and how rich is their feedback in each of these e-feedback media?

A statistical comparison of the tutor’s use of feedback in the four media showed similar patterns of use in the feedback summary forms related to written and spoken assignments. In both cases, the proportion of the tutor’s comments relating to strengths was higher than the proportion of comments related to weaknesses. Comments on the content and structure of the assignment also occurred in a higher proportion than comments on the use of language. In terms of depth, feedback tended to categorise strengths and weaknesses more often than simply identifying them, or giving corrections, explanations or advice for future assignments (see Table 3). All these patterns were even more pronounced in the summary forms related to written assignments.

Table 3. Analysing assignment feedback using the FACT tool: examples from Spanish language assignments.

<table>
<thead>
<tr>
<th>Depth</th>
<th>Focus on weakness</th>
<th>Focus on strength</th>
<th>Depth</th>
</tr>
</thead>
<tbody>
<tr>
<td>Error identified only</td>
<td>Manos blancones</td>
<td>Good work</td>
<td>Strength identified only</td>
</tr>
<tr>
<td>Error categorised, but not corrected</td>
<td>Manos blancones → Agreement</td>
<td>You use a wide range of language structures</td>
<td>Strengths categorised or described as per marking criteria</td>
</tr>
<tr>
<td></td>
<td>No critical evaluation</td>
<td>You identified the most relevant key issues</td>
<td></td>
</tr>
<tr>
<td>Error corrected</td>
<td>Manos blancones → blancones</td>
<td>No digo que quieran... → Good use of the subjunctive</td>
<td>Illustrated with specific example from student’s performance</td>
</tr>
<tr>
<td></td>
<td>...for example you could have mentioned that global section has many distractors</td>
<td>Your comparison to Colombia is particularly sound</td>
<td></td>
</tr>
<tr>
<td>Explanation given</td>
<td>Although it ends in O, ‘mano’ is a feminine noun</td>
<td>This connector makes it very clear that a new section is starting here.</td>
<td>Explanation given</td>
</tr>
<tr>
<td></td>
<td>You need to show that you can look at the facts from different perspectives</td>
<td>This reference would enable any reader to find your original source easily.</td>
<td></td>
</tr>
<tr>
<td>Explicit advice given for future performance</td>
<td>Revise section 6.1 of your grammar book</td>
<td>Keep using this approach as it is very effective</td>
<td>Explicit advice given for future performance (not “you could/should have”, but “try also...”)</td>
</tr>
<tr>
<td></td>
<td>Get into the habit of analysing the question word by word, this will help you ensure that your answer is relevant.</td>
<td>You may also find it interesting to look up the following reference...</td>
<td></td>
</tr>
</tbody>
</table>

Written annotations and audio feedback displayed the opposite tendency. In both of these media, the proportion of the tutor’s comments relating to weaknesses was higher than the proportion of comments related to strengths. Errors were often addressed at more than one depth (e.g. explained or categorized as well as corrected). Comments
on use of language were found in a higher proportion than those relating to content and structure. In terms of depth, a higher proportion of corrections was found in these two media. There were also more explanations and fewer categorisations of strengths and weaknesses in these media. The higher proportion of explanations was most pronounced in the audio feedback, which suggests that the medium lends itself particularly well to specific explanations on selected aspects of a student’s performance. Indeed the number of items discussed by tutors in their audio feedback was generally smaller than the number of items covered in other media.

An analysis of the total number of comments in the sample (including annotations on scripts) shows the vast majority of tutor’s comments relate to language errors, and that feedback on language use focuses more on weaknesses than does feedback on content (see Figure 7).

Figure 7. Total number of comments on strengths and weaknesses relating to content and language.

In terms of overall distribution, feedback related to weaknesses tends to be deeper than feedback related to strengths (see Figure 8).

Figure 8. Percentage of comments on strengths and weaknesses according to depth.

Feedback related to the use of language is also deeper than feedback related to the content of the assignment (see Figure 9).
4.3 Student engagement with the feedback

How do students engage with the feedback given by tutors on their assignments?

Due to the nature of the feedback on feedback exercise (screencast recordings), only the students’ engagement with written feedback could be evaluated. A comparison between their responses to the summary form and to annotations on the script might have been possible, but for reasons of time our priority was to evaluate their responses to different depths in the feedback on the strengths and weaknesses of their work. Data analysis shows that highly motivated students do engage with tutor feedback and make active efforts to integrate it. However in some cases their cognitive, affective, or metacognitive responses to the feedback are ineffective. This suggests that a tutor’s incorrect assumptions about the student’s abilities, expectations or attitudes in relation to feedback can contribute to these occasional breakdowns in communication.

The participants in this study were highly motivated students, and therefore it would be reasonable to expect a considerable number of responses indicating that effective feedback dialogue was taking place. Indeed, cognitive responses to feedback on weaknesses, especially those related to what students regarded as “silly mistakes” (spelling, agreement, missing references, etc.), tended to result in the construction of knowledge through active integration or informed acceptance of the tutor’s feedback. Positive affective responses to feedback on strengths were also very common, especially when students perceived that their efforts were being acknowledged or the tutor was establishing a personal rapport with them (“she spotted I am French, well done tutor”). Metacognitive responses to feedback on strengths in the form of planning strategies to improve future performance were also common.

However, somewhat unexpectedly in a group as motivated and high-achieving as this, a number of responses indicating ineffective feedback dialogue were also found alongside these constructive responses. When feedback focused on errors relating to more complex structures, for example when a correction to the syntax of a sentence was left unexplained, the students’ cognitive responses were ineffective as they either
Accepted the feedback without understanding it, or failed in their attempts to integrate it.

At more advanced levels, unexplained lexical corrections were perceived by students as the tutor’s failure to appreciate their creative attempts at experimenting with the language through the use of metaphors. This caused them to reject the feedback both on cognitive and affective grounds, as they felt that their personal efforts had not been appreciated. Well-intended tutor support was also rejected when students suspected a one-size-fits-all approach that failed to take their individuality into account (e.g. lengthy technical tips given to a student who had worked for years in IT, cut-and-paste invitation to contact the tutor at the end of a feedback form, etc.).

The presence in the sample of responses indicating both effective and ineffective feedback dialogue is consistent with claims commonly voiced by tutors that their feedback, or at least some of it, often does not achieve its intended purpose. The roots of the communication breakdown may be cognitive, as in cases where the depth of feedback was not sufficient, or affective as when students felt that their efforts or individuality were not being duly acknowledged. The fact that even a highly motivated group of students such as the participants in this study occasionally failed to integrate some of the tutor feedback suggests that this type of occurrence might be considerably more common in a sample including a wider range of abilities and motivational levels.

Peer-to-peer feedback was also used as a means of evaluating to what degree language students engage with their tutor’s feedback. A preliminary comparison between the type of feedback produced by tutors and their students shows similarities between their respective approaches. Results show that students are able to reproduce and reuse feedback-specific vocabulary in an appropriate and context-specific manner. This suggests that students have indeed engaged with the feedback received from their tutors and are therefore able to consider similar aspects of performance in the same way when it comes to giving feedback themselves.

5 Conclusions and Recommendations
Overall, e-Feedback was seen as an improvement over conventional feedback by most language tutors surveyed, both in terms of quality of the students’ learning experience and of the tutors’ teaching experience. This is relevant to the Open University, which is a major stakeholder in its introduction.

OU responses regarding the perceived usefulness of different media indicate that both students and tutors do not currently value audio feedback as much as written media. On the other hand, audio feedback has been found to elicit a greater proportion of explanations. Furthermore, its use as a means of giving feedback on written assignments was seen as very valuable by UoM students, which suggests that the OU may not be making the most of the medium’s affordances. In their audio feedback OU language tutors are currently advised to concentrate on correcting pronunciation and intonation, but they may be missing an opportunity to give more feedback (especially explanations) on other aspects of performance.
The FACT analysis of over 200 assignments revealed a consistent tendency to give deeper feedback on weaknesses than on strengths. Feedback related to the use of language also tends to be significantly deeper than feedback related to the content of the assignment, and feedback on language use focuses more on weaknesses than does feedback on content. Although one might be tempted to advise that tutors try to reduce such differences, these may well respond to a genuine need and further reflection must be undertaken before drawing a set of specific recommendations in one particular direction or another. Nevertheless, the patterns must be noted and critically evaluated by the stakeholders. A working group comprising members of the eFEP team, staff tutors and central academic staff has been put in place in order to identify which of the observed patterns are justified, which ones may need to change, and how the briefing and training of OU tutors and monitors can be improved in the light of this evaluation. Further data on student engagement should also give us a better idea of which aspects of feedback need to be addressed.

Feedback on feedback proved to be a powerful evaluation tool for that purpose as we learnt that even highly motivated high achievers sometimes engage in unhelpful strategies as a result of missing or insufficient explanations in the feedback. Some students in the advanced level even displayed negative emotional responses as they perceived certain corrections as a failure on their tutor’s part to appreciate their creative use of language. By making such misunderstandings apparent, feedback on feedback constitutes an excellent means of improving feedback alignment and fruitful dialogue between students and tutors.

Project findings clearly demonstrate the potential value of feedback on feedback and peer-to-peer feedback for promoting learner engagement with assignment feedback, and we recommend that they are both integrated at some point into the curriculum. Tutors may also use them at their discretion, possibly at the beginning of a module for diagnostic purposes or when they feel that a student does not appear to learn from the feedback provided.

6 Lessons learned
This was an ambitious project encompassing a wide range of evaluation approaches. Such triangulation was extremely productive and yielded interesting results. However the different strands were at times difficult to manage with only a small team of three. The size of the team made us vulnerable to risk factors related to staffing, as indeed we found when a member of the team had to take sick leave for several months. Despite these challenges, we now have an extremely rich dataset that will generate further research over the next couple of years, and could be at the start a follow-up project, e.g. qualitative analysis of peer-to-peer feedback using NVivo, development of an online evaluation tool for feedback alignment, etc.

We were fortunate to have access to the expertise of the Open University’s Institute of Educational Technology (IET), who provided advice on sample selection and questionnaire design, developed the online version of the questionnaires, managed their delivery and conducted basic descriptive analysis.
Appointing consultants with the right skills proved to be more of a challenge, although once appointed, they were extremely helpful and competent. FACT coding and its related statistical analysis would not have been possible without their help.

As a method, FACT was initially conceived as a means of generating clearly quantified profiles of tutors’ feedback for training purposes. However, quantifying the comments in a given category proved to be an extremely complex process, only worth undertaking within a research setup. It took three rounds of revisions to reach the desired level of inter-rater reliability, although the final version produced a 0.97 IR index. From a training perspective, the value of FACT proved to be as a stimulus for reflection rather than visual evidence of individual profiles, as originally envisaged. A simplified resource was subsequently produced in the form of a self-reflection checklist based on the FACT criteria.

A second resource was also produced in the form of a set of templates for a feedback alignment exercise in which students and tutors are guided through a parallel analysis of the main strengths and weaknesses in a particular assignment and compare notes at the end. This approach, however, was not as popular and none of the tutors to whom it was presented took up the suggestion to try it out for themselves. This could be because the layout of the printed material made it seem more time-consuming than it really was. The method is probably better suited to online delivery and we are discussing the development of a prototype as a possible follow-up to the project.

Feedback on feedback, on the other hand, generated considerable interest at both internal and external events (e.g. the paper in which it was presented received the EDEN 2013 Best Research Paper Award in Oslo), and we also anticipate much interest in our planned follow-up of peer-to-peer evaluation.

Teaching institutions in the UK and beyond which have expressed interest in following up on the findings of the eFEP project are as follows:

- We were approached by Dr. I-Chant Chiang (Aberystwyth University) regarding a possible collaboration on a follow-up to the project Optimising Audio Feedback Assisted Learning for Student and Staff Experience and agreed to act as critical friends for this follow-up project. Unfortunately funding was not granted in this occasion, but we shall resume contact in September 2013 in order to explore further opportunities for collaboration.

- We were also invited in May 2013 to give a workshop at the Business School, University of Hull, which is currently reviewing its assessment and feedback strategy.

- We are currently working on a proposal for a follow-up project in collaboration with Denise Whitelock (Open University) in order to develop an electronic version of the feedback alignment exercise.
• Dr Jos Tolboom, from the Netherlands Institute for Curriculum Development, approached us via the project website offering to contribute to this new development. His subject area is in Maths and Statistics.

• Dr Cathie Harrison, from the Australian Catholic University (Early Childhood Education), who attended our presentation at EDEN 2013, is also interested in collaborating with us in the area of feedback on feedback.

Potentially, these new contacts could broaden the scope of the project beyond Languages to other subject areas including to Business Studies, Maths and Statistics, and Education, and beyond the UK.

7 References