# An Automated Individual Feedback and Marking System: An Empirical Study

**Trevor Barker**
**University of Hertfordshire, Hatfield, UK**
t.1.barker@herts.ac.uk

**Abstract**: The recent National Students Survey showed that feedback to students was an ongoing problem in Higher Education. This paper reports on the extension of our past research into the provision of automated feedback for objective testing. In the research presented here, the system has been further developed for marking practical and essay questions and providing automated feedback. Recent research at the University of Hertfordshire was able to show that learners and tutors accept and value our automated feedback approach based on objective tests and Computer Adaptive Testing. The research reported in this paper is an important extension to this work. The automated feedback system developed for objective testing has been extended to include practical testing and essay type questions. The automated feedback system, which can be used within any subject area, is based on a simple marking scheme created by the subject tutor as a text file according to a simple template. Marks for each option and a set of feedback statements are held within a database on a computer. As marks are awarded for each question by the teacher an individual feedback file is created automatically for each learner. Teachers may also add and modify comments to each learner and save additional feedback to the database for later use. Each individual feedback file was emailed automatically to learners. The development of the system is explained in the paper and testing and evaluation with 350 first year (1 final practical test), 120 second year (1 written and 1 practical tests) and 100 final year (1 final practical test) undergraduate Computer Science students is reported. It was found that the time to mark practical and essay type tests was reduced by more than 30% in all cases compared to previous years. More importantly it was possible to provide good quality individual feedback to learners rapidly. Feedback was delivered to all within three weeks of the test submission date. In end of module tests it was very beneficial indeed as it had proven difficult to provide feedback in the past after modules had ended. Examples of the feedback provided are presented in the paper and the development of the system using a user-centred approach based on student and staff evaluation is explained. The comments of staff teaching on these modules and a sample of students who took part in this series of evaluations of the system are presented. The results of these evaluations were very positive and are reported in the paper, showing the changes that were made to the system at each iteration of the development cycle. The provision of fast effective feedback is vital and this system was found to be an important addition to the tools available.

**Keywords**: assessment, feedback, automated systems, development, evaluation

## 1. Introduction

High staff/student ratios often mean that tutors often have great difficulty in providing students with high quality feedback on assessment performance that is timely and meaningful. This is despite the many advances in computer aided assessment and technology. Chickering and Gamson (1987) list 'prompt feedback' amongst their seven recommendations for good practice in teaching. Promptness is an important factor but other factors are equally so. For example feedback must be constructive, appropriate, useful, accurate, individual, delivered in context, detailed and should also facilitate feed-forward. Freeman & Lewis (1998) amongst others have reported on the importance of feedback as a motivator for student learning. Given the increasing pressures on teachers' time, these goals are becoming increasingly difficult to achieve. Thus, there is an increasing demand for the development of software applications that would enable the provision of timely, individual and meaningful feedback to those learners. In previous work the author and colleagues have reported on the use of automated feedback systems related to Computer Adaptive Testing (CAT) (Barker, 2009; Lilley et al, 2004). In this research it was found that the systems developed had many of the benefits listed above. Constructive detailed feedback was delivered quickly, accurately, in context and it was possible to facilitate appropriate feed-forward for individual learners. In addition feedback on the cognitive level at which learners were effectively working was provided, related to Bloom's taxonomy (Bloom, 1956; Anderson & Krathwohl, 2001). CAT systems were tested and evaluated by staff and students and shown to be effective and highly valued (Lilley et al., 2004). For feedback to be effective, it is argued, it should be individual for each learner and timely. The use of our feedback system based on objective computer-based adaptive testing was shown to be effective, but limited as it could only be applied to CATs. The CAT applications that were used in our feedback test systems have been reported by Lilley and colleagues (Lilley & Barker, 2002; 2003; 2004; Lilley et al., 2004; 2005). However it is important that fast and effective feedback be provided for a wider range of tests, especially practical tests which occupy a significant amount of the assessments in the domain of Computer Science and

also essay and written text questions. These are not catered for at all in CATs. The work described in this paper therefore relates to the development, testing and use of an automated marking and feedback system for essay and practical assessment.

### 1.1 Feedback provision

The use of technology in order to support learning has been shown to be highly regarded and expected by learners (Parkin & Thorpe, 2009). There is evidence from previous research in the literature that many students expect to receive their grades and feedback online using the affordances that technology brings to learning (Bloxham & Boyd, 2007). Hepplestone and Mather (2007) provide supporting evidence for the importance of providing online feedback to learners via their 'feedback wizard'. Students value the flexibility, privacy and convenience of receiving feedback in this way. Price & O'Donovan (2008) suggest that providing feedback in this way engages students and allows them to respond to feedback at a time when they are emotionally prepared to do so. The timeliness of feedback has also been stressed in the literature, for example Mutch (2003) emphasizes the importance of providing feedback at a time and in a context when it is still meaningful to learners. This is particularly important in the provision of feed-forward in order to guide future learning and preparation for future assignments. Winter & Dye (2004) have shown that students are less likely to collect their feedback and grades unless they are provided within a reasonable time from the assignment. The format of feedback has also been shown to be important in studies. Feedback that is typed, and in a clear and legible format has been shown to be more acceptable to learners than other forms, for example hand written feedback (Bridge & Appleyard, 2005; Denton et al, 2008). Feedback system that deliver well formatted text via an online system were likely to be more highly valued by learners. Automated systems would also be expected to be more efficient and reduce the time taken by tutors to mark work (Jones & Behrens, 2003). Based on the evidence from the literature summarized here, it was decided to develop an automated feedback system that would deliver typed and formatted feedback to learners in a fast and secure fashion using electronic mail.

### 1.2 Methodology

The rapid development of computer software systems is readily facilitated using a user-centred prototyping approach (Sommerville, 2010). Prototyping has several benefits according to Sommerville These include the clarification of requirements, providing a focus and direction for designing the system, and importantly the involvement of the stakeholders in the project and end users of the system (Sommerville, 2010). It also allows the system developer insight into the accuracy of initial project objectives and whether they can be successfully met. In complex domains such as those found in education, a user-centred prototyping approach is essential in order to evaluate the effectiveness of the system in a full context (Barker & Barker, 2002). In the development of the automated feedback system described here, expert users (tutors and assignment moderators) and small groups of end users (student recipients of the feedback) were employed in order to guide the prototype iterations. Nielsen and Mack (1994) have shown that small groups of expert users are able to facilitate rapid and efficient evaluation of computer systems in this way in order to support rapid prototyping.

## 2. Requirements of the system

In complex domains such as teaching and learning, the evaluation of implemented systems by stakeholders at all stage of the development process is absolutely vital as explained by Barker and Barker (2002). Teachers and learners as well as other stake-holders were expected to have a significant input into the nature of the system developed. The first stage of this was to develop a set of requirements for the system that would enable implementation, testing and subsequent improvements to the system. The initial requirements of the proposed system therefore, were arrived at as follows. Based on a survey of the literature summarized above and the evaluation of previous automated feedback system (Lilley et al., 2004) which included significant input from staff, students and academic managers a list of desirable functional requirements for the proposed system was produced. After discussion with colleagues and modification a ***basic*** set of functions was produced for the design of a first stage prototype. These are presented in table one. The ten functions shown in table one, were considered to be the *minimum* set necessary in order to develop the first stage prototype. This prototype would be developed and the implementation tested and evaluated in a real context. It was intended then that the results of this evaluation would enable the production of an improved set of requirements based on this experience.

 

**Table 1**: List of agreed functional requirements for the feedback system

| Function |
|---|
| The system should be a computer-based marking system. |
| Simple to install and useful for a range of assessments and assessment types. |
| It should be able to mark both practical and essay type questions. |
| It should provide fast feedback. |
| The list of students and email addresses to be read in from university admin system in order to minimize work for the teacher. |
| Teachers would be able to enter five levels of feedback for each question. |
| General feedback would be allocated for each question based on the mark awarded. |
| The system would collate marks and produce feedback records for teachers. |
| Individual feedback and marks for each learner to be saved to a database file. |
| Feedback and marks to be distributed via electronic mail after checking. |

## 3. Development cycles

The first prototype was developed using a standard Microsoft event driven programming language. This was decided upon mostly for speed and for ease of installation and testing. The system consisted of three main parts. A feedback file that contained the general feedback for each question, a student file that contained the list of students and their details, provided by the university admin system and a graphical user interface that read in the feedback and student files in order to allocate marks and feedback. The output from the system was a file which contained marks and feedback suitable for distribution via electronic mail. This was achieved by using a simple mail merge application within a Microsoft word processing application that read the file and applied it to a mail merge template developed for this purpose. Figure one shows the first prototype developed in this study.

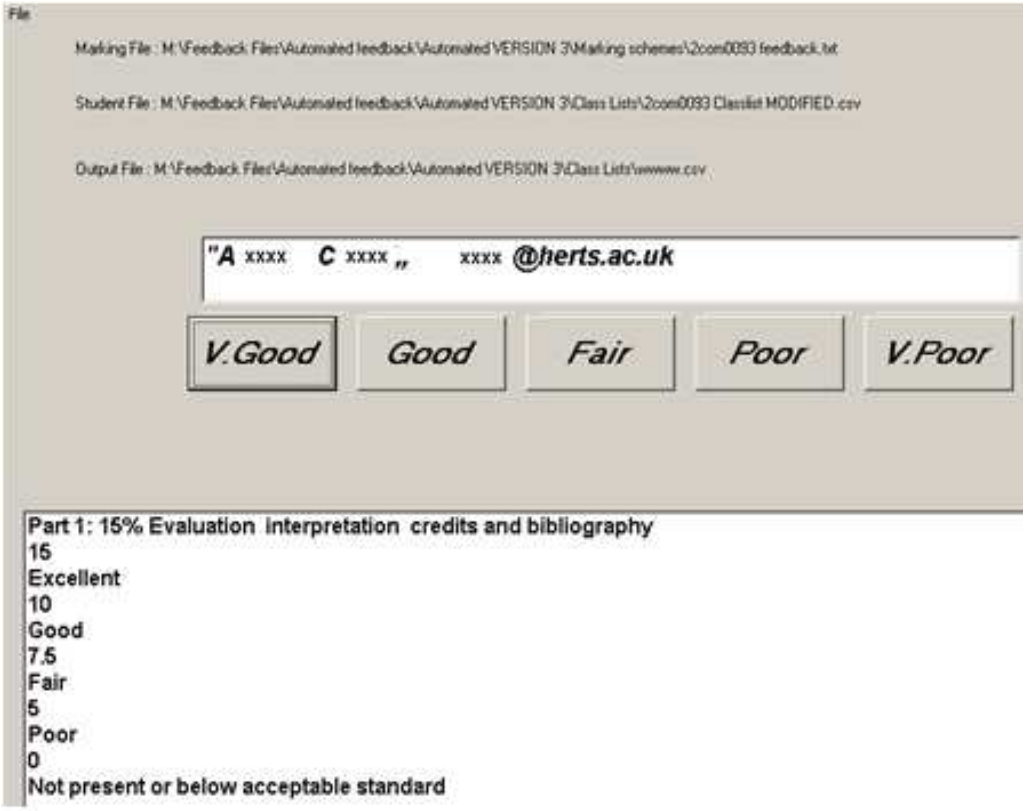


**Figure 1**: First version of the automated feedback and marking system

The prototype shown in figure one above was used to mark a summative practical test for a set of approximately 350 first year computer science students. This test was taken under supervised conditions in a computer laboratory. Practical work was uploaded to the University's managed learning environment (MLE) for marking. The test consisted of 16 questions. The text box below the buttons presents performance indicators for the mark. The buttons are used to allocate feedback for each question and the actual mark awarded for each section was entered by the marker after the appropriate button had been selected. An example of the feedback provided for each mark range is shown in table 2 below.

### 3.1 Format of the feedback

The feedback file created for use in the prototype was developed based upon a marking scheme for the assignment. For each question in the assignment five general feedback statements were written for excellent, good, fair, poor and absent or below acceptable standards. After discussion between markers these were manually written into file which suitable for reading by the prototype. Table two below shows an example of the feedback provided for one question

**Table 2**: Example of feedback range provided for one question

| Part 4: 10% Sequence of still photographs |
|---|
| 10: Excellent sequence of stills with excellent subject matter and high quality images |
| 7.5: Good sequence of stills with good subject matter and good quality images |
| 5: Fair sequence of stills with some issues with either subject matter or the quality of images |
| 2.5: Poor sequence of stills with considerable issues with either subject matter or the quality of images |
| 0: Sequence of stills was not present or below acceptable standard |

The number in each section is a guide to the marks relating to each of the feedback comments.

### 3.2 Use of the prototype

The first stage prototype shown in figure one above was used to mark approximately 350 practical assignments over a one week period. After the marking was completed marks were transferred to a spreadsheet in order to check that no errors had been made in marking, markers were consistent and that the mean and other statistical measures for the test were similar to other tests on the module. A sample of marked work and feedback was then passed to an external marker to be moderated and his comments were received for later analysis.

Once the course team was satisfied that the test had been marked fairly and accurately, marks and feedback were released via electronic mail to individual learners. In previous years it had proven difficult to achieve this timescale with smaller groups of approximately 250 students within a six week period. On this occasion we were able to release the marks three weeks after the end of the assignment for a group of 350 learners. Markers reported that the marking itself was faster and more efficient, taking approximately 30% less time to mark the work than previously. The greatest saving on time was related to writing and distributing feedback. On some occasions in the past it had not been possible to deliver feedback until after the end on the course itself and on one occasion feedback was not delivered at all since students were on their summer vacation by the time feedback was ready for distribution.

## 4. Evaluation of the first prototype

Approximately one week after the marks and feedback had been distributed, markers met to discuss and reflect upon the exercise. Comment from the external marker were also distributed and considered. Fifteen students were selected quasi-randomly to answer a short questionnaire. Selection was based on their scores obtained in the test. It was important that students with a range of scores in the test had an opportunity to comment on the feedback provided by the system, so five students were selected in each on the performance ranges, under 50, between 50 and 75 and above 75 marks. Table three presents a summary of their responses to the questionnaire.

 

**Table 3**: Learners' responses to questionnaire and score achieved in the test

| Likert Scale (1 to 5)<br>1=disagree, 2, somewhat disagree, 3=neither agree or disagree, 4 somewhat agree, 5 = agree | Student responses (n=15) | | |
|---|---|---|---|
| Statement | score > 75 | score 50-75 | score < 50 |
| Feedback was useful to me | 4.0 | 3.8 | 3.2 |
| Feedback was fast | 4.8 | 4.6 | 4.6 |
| Feedback was delivered conveniently | 4.8 | 4.2 | 4.4 |
| Feedback was fair | 4.4 | 3.8 | 3.8 |
| The amount of feedback was good | 4.2 | 4.0 | 4.2 |

Table three suggests that students felt that the feedback was in general timely, fair and useful and delivered in an acceptable manner. Inspection of table 4 shows that in general, attitude to feedback did seem to depend on the mark achieved by learners on the test, though in general attitude was good for all categories. Markers considered that the marking and feedback system was a good idea and a valuable tool to help in the rapid delivery of marks and feedback. All agreed that it had operated flawlessly and that marking time had been reduced considerably. In all only 9 of the 350 students reported any problems with the marks awarded to them, none of which related to the performance of the prototype itself.

Several issues however were raised by markers and the external moderator relating to the feedback and functions available in first stage prototype. Perhaps the most important related to the flexibility of the feedback provided. Although markers considered it extremely useful it was considered to be very inflexible. Feedback comments of the type shown in table two above were considered to be too general and inflexible by the markers. For example in table two the feedback statement *"Excellent sequence of stills with excellent subject matter and high quality images*" relates to image quality and subject matter. It would be an improvement to separate this into two sections and provide a mark and appropriate feedback for each

Markers also wanted to add their own feedback comments on the assignment related to aspects of the performance overall. It was also considered useful to include some feedback related to the completeness of the work handed in. These two features were added to the list of functional requirements presented in table one to be used in the development of the next stage prototype.

In order to make the feedback more relevant, marking schemes were re-written in such a way that feedback could be related more specifically to performance. This was achieved by breaking each of the questions into smaller parts and writing the marking scheme and feedback comments to reflect this more closely. In this way a larger amount of more directed feedback could be written relating to each section of a question.

Another suggested improvement was the replacement of the rather inefficient way in which marks and feedback were allocated using the buttons shown in figure one. It was suggested that marks be awarded each section of a question and that feedback would be presented based on the mark awarded. In the previous system feedback was awarded by selecting the appropriate button and the mark entered later. This modification would do away with the buttons altogether and make the use of the system more efficient. It was also decided to produce additional introduction and summary screens to show in the first place the submission requirements for each assignment and also a final screen summarizing the marks and feedback for each learner.

## 5. Development of the second prototype

The modifications were made to the system as outlined in section 4 above and a second stage prototype of the system was developed and tested prior to use with students on an assignment. The modified version of the system allowed markers to comment on the completeness of the hand-in for the assignment as shown in figure two. In this version, the hand-in information is presented to the marker who may then make additional comments on the completeness or nature of the hand-in.

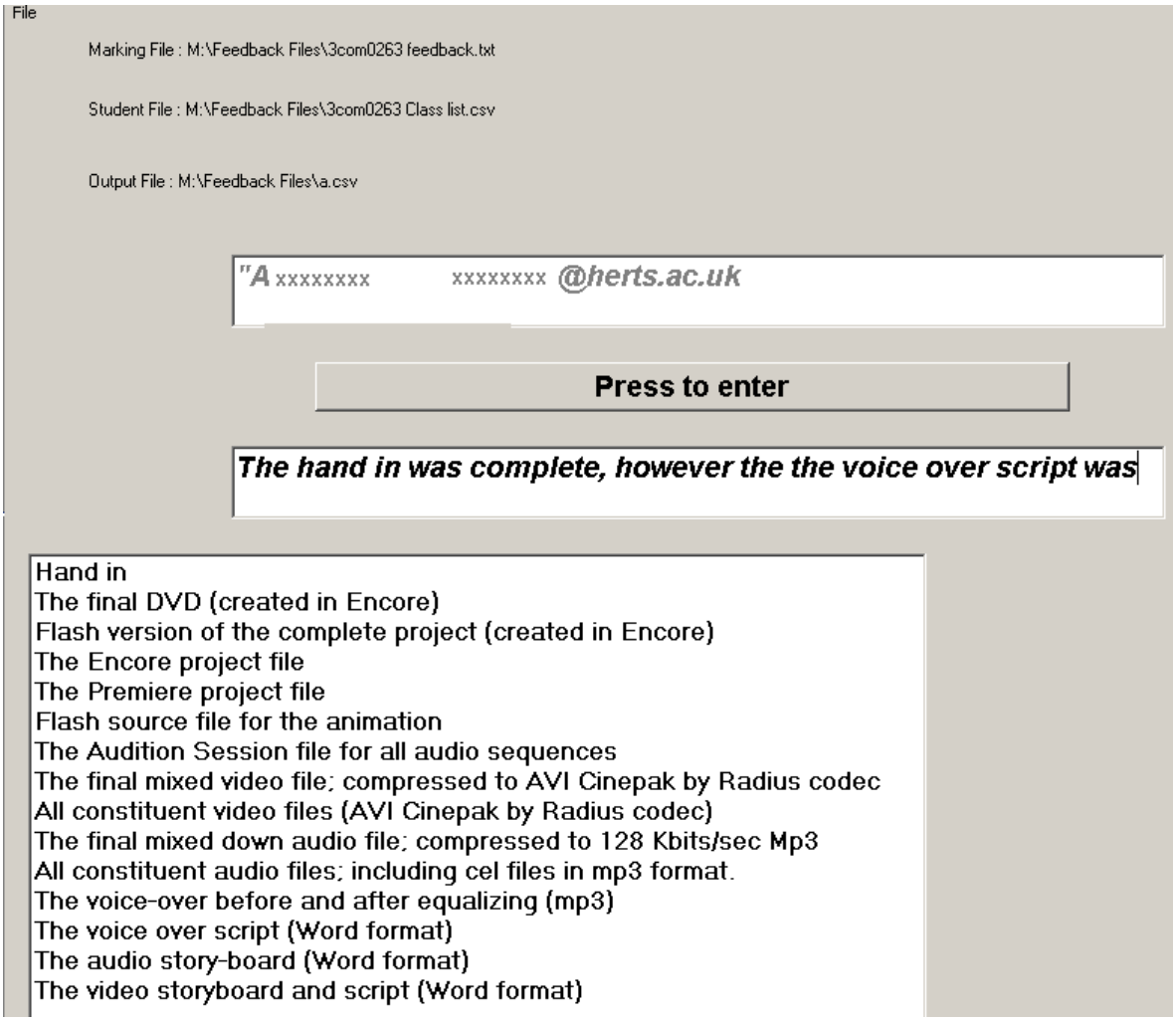


**Figure 2**: Modified version of the system, showing hand-in information and tutor-entered feedback

Each of the questions and question parts were presented by the system and the marker was simply required to enter a mark. Appropriate automated feedback was determined by the system based on the mark awarded in each section of a question, reading it from the database file for the assignment. After all the question sections had been marked, the system presented a final summary screen so that the marker could check that the marks had been awarded accurately. If this were not the case, the marker could cancel the entry and mark the student again. An example of the summary screen is shown in figure three.

## 6. Evaluation of the second prototype

The second prototype was used to mark three assignments, a second year practical summative assessment for 125 students and two final year summative assessments for approximately 100 students. These assignments were complex in that they contained both practical and theoretical elements and were completed over a six week period. In the past it had been extremely difficult to achieve consistent marking in this type of assignment. They were slow to mark and feedback delivered in terms of quality and quantity was fairly inconsistent between markers.

 

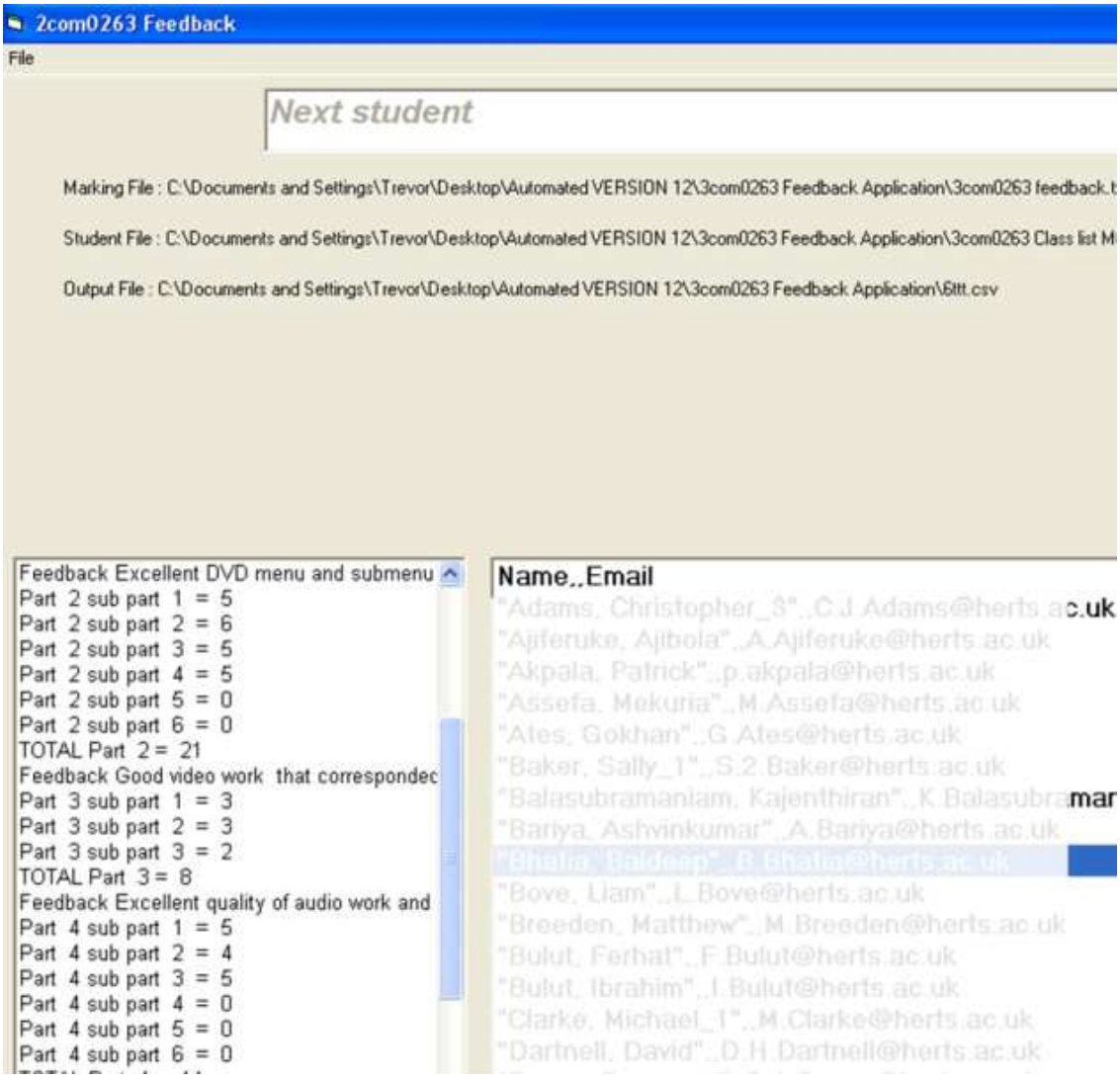


**Figure 3:** Modified version of the prototype system, showing final summary screen

After the assignments had been marked, moderated and subject to quality measures, marks and feedback were released as before. In this case, a sample of 15 students only from the second year module was selected using the same sampling method as before in order to test student attitude to the feedback using the same questionnaire as before. The results of this are shown in table four below.

**Table 4**: Second year learners' responses to questionnaire and scores achieved in the test

| (Likert Scale (1 to 5)<br>1=disagree, 2, somewhat disagree, 3=neither agree or disagree, 4 somewhat agree, 5 = agree | Student responses (n=15) | | |
|---|---|---|---|
| Statement | score > 75 | score 50-75 | score < 50 |
| Feedback was useful to me | 4.8 | 4.2 | 4.0 |
| Feedback was fast | 4.8 | 5 | 4.6 |
| Feedback was delivered conveniently | 5 | 4.8 | 4.8 |
| Feedback was fair | 4.6 | 4.4 | 4.4 |
| The amount of feedback was good | 4.8 | 4.6 | 4.4 |

As before, the attitude of learners to the feedback was positive. The small sample size (n=15) was chosen in order to facilitate rapid development of the prototype rather than to identify significant differences between the means displayed in table 4. It was therefore not possible to perform a statistical analysis on these results. However cautious comparison of the results displayed in table 4 suggests that learners were slightly more satisfied with modified prototype than with the earlier version. A larger study to investigate the significance of any differences in attitude related to performance will be undertaken when the final version of the prototype becomes available.

Tutors were also invited to discuss their experiences of using the second prototype system and these were taken along with comments from the external marker. It was generally agreed that the system was improved by the modifications in terms of efficiency of marking and the quality of feedback presented to learners. As before the system performed faultlessly and there were no problems with installation or the automated distribution of feedback via electronic mail. One issue that surface is that although learners were in general satisfied with the feedback awarded, there was a greater number of students prepared to challenge their mark. It is suggested that this was due to the more detailed and specific marking scheme provided along with more detailed feedback relating specifically to each question part. Errors in marking made by markers were more readily identified by students and these were naturally more likely to be questioned. This issue was considered to be a positive feature by markers of the system, leading to greater accuracy and fairness of marking. The tutor-entered feedback on the hand-in and the general comments were considered to be a good feature of the modified version. The external marker considered the system to be extremely useful and made several highly supportive comments related to the system. The quality and quantity of feedback provided was considered to be much batter in the second prototype than in the first.

### 6.1 Suggested improvements

Based on the findings from the evaluation of the second stage prototype it was decided to produce a third version with new modifications suggested by the markers and external moderator. It was suggested that tutors should be able to add to or modify the automated feedback at each stage of marking as well as adding general comments at the end. It was further suggested that these additional comments be saved to a database file so that they might be used again later. In addition an option to show an image of the student was suggested, although this suggestion did not receive universal support from all present as some favoured more anonymous marking.

## 7. Development of the third prototype

In order to allow for the comments from markers and the external moderator a third stage prototype is currently being developed. In this version several improvements are being made.

A more robust and secure database system is being employed which will not only store the automated feedback for each question section but will also store additional feedback comments that are added to each question to supplement the automated feedback. A list of these is then available when the question is again marked for another learner and can be added to the feedback with a single click.

Developing a high quality feedback database is an important part of the system and at present this is a time consuming and demanding task. In order to simplify this task, the system is being modified in order to make it simpler for the tutor to enter their marking scheme and feedback statements and create a feedback file that can be used by the system. At present the teacher must create a text file containing a great deal of meta-information relating to the format of the questions along with the feedback for each option. In the new system the tutor need only enter the marking scheme and the feedback. The feedback database file is then formatted by the system.

An option to display an image of each student has also been added, although this may be switched on or off. It was also suggested that it be possible to modify marks without needing to re-enter the complete set of marks for a student and this feature was added. Finally the feedback distribution has been improved so that it is no longer necessary to use a word processor to distribute the feedback via electronic mail merge. This function is now achieved from within the application directly. Figure four shows an example screen of the latest prototype. In this version the user interface has been improved. Teachers are able to set up assessments and feedback much more easily.

 

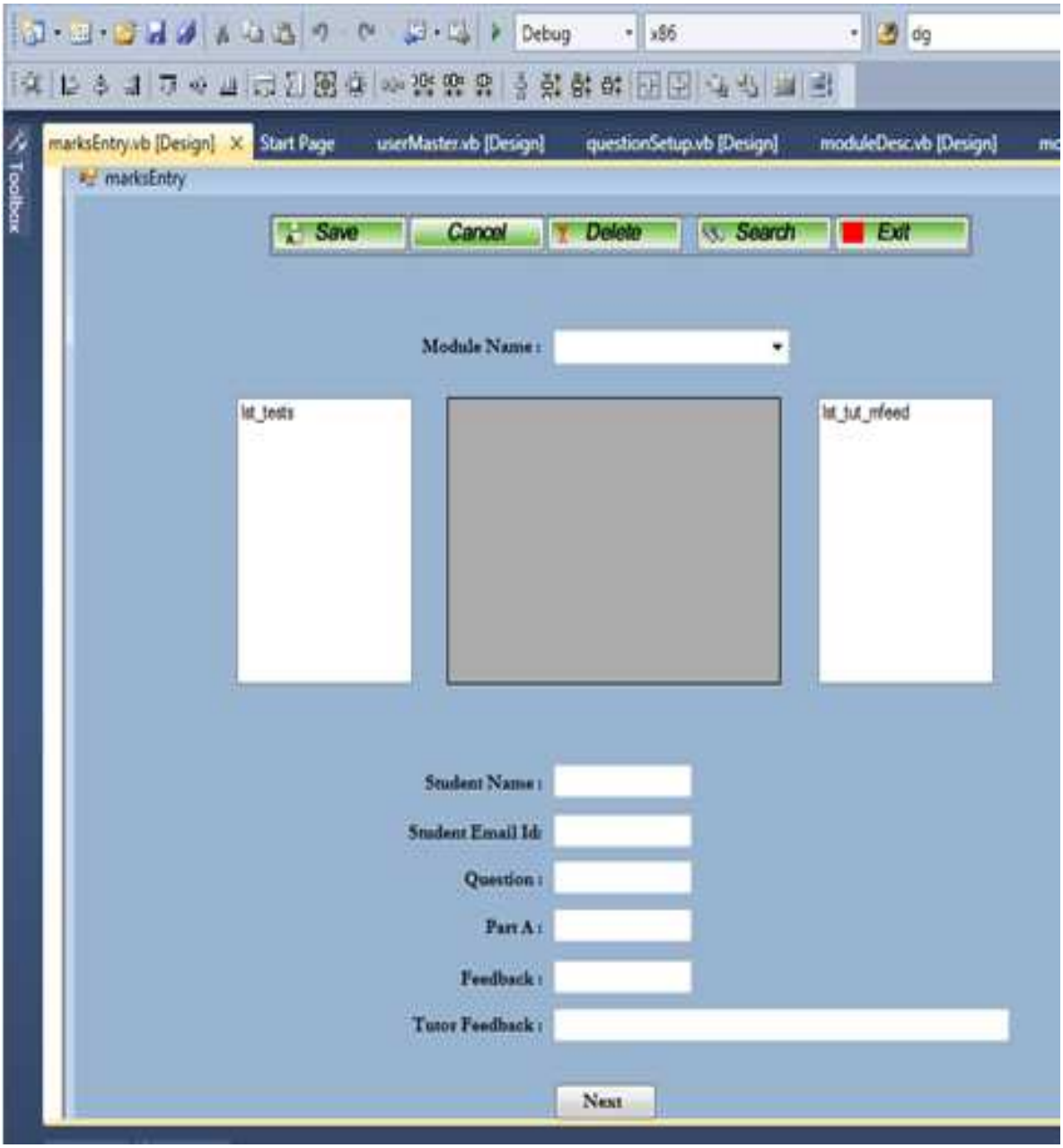


**Figure 4**: The latest version of the prototype showing the modified user interface

The systems described in the previous sections were used to provide feedback for more than 800 students. Table five shows the breakdown of the students, levels and modules for which the application was employed.

**Table 5**: Students, levels and modules and assignment details for which the systems were used

| Module | Number | Level | Assignment type |
|---|---|---|---|
| E Media Design | 320 | First year BSc | Practical |
| Games Development | 180 | Second year BSc | Practical and theory |
| New Entertainment systems | 160 | Third Year BSc | Practical and theory |
| Multimedia System Design | 140 | MSc | Theory |
| Referred Deferred coursework | 50 | All levels | Practical and theory |

The software was used for a fairly wide range of learners, from first year through to master's level for a range of practical and theory assignments. More than 850 students received feedback from the system. Data from student feedback questionnaires suggested that the feedback delivered was of good quality. Data was only available for approximately half of the modules involved in the study. This data took the form of written comments. In general comments were complimentary about the system. Of 24 comments related to the feedback provided only 1 was slightly negative.

> *"Although the feedback was quick I thought I did better than the grade I got". (Student level 1)*

This was a general problem. The more detailed the feedback provided the more likely students were to question their grades. Most students provided a more positive view of the feedback they were provided with as the following example shows.

> *"… feedback was really fast and delivery by email is a great idea. I got this 3 weeks after. On other modules we get marks months after hand in". (Student level 3).*

The majority of comments from students were similar to that of the level 3 student above. In general it suggested that the application functioned well and the feedback provided was good and appropriate.

## 8. Use of the modified system

Prototype version four has been developed, however it has experienced several problems and despite having additional functionality it is not robust enough for reliable use. Further development work is required on this system before it can be used generally. However, based on the positive results obtained from the evaluation of the use of the third prototype, t was decided to extend the use of this system to include more modules at BSc and MSc level and hence more learners. In all, more than one thousand additional assessments have been undertaken with the third iteration prototype so far this year. An important addition was the use of the prototype to assess group project work at master's level. Previously such work had proven difficulty to assess and difficult to provide learners with good and timely feedback because of the complexity of marking.

The module had four summative assignments. The first was an individual online multiple choice test covering the principles of multimedia design. The work leading up to the first assignment was intended to prepare learners for the latter three assignments. The 55 student on the module then split into 22 groups of between 2 and 4 persons. The second group assignment related to the development of a prototype Flash website in which students, in groups, produce a minimal content software prototype: essentially the basic structure of the website together with an animation to promote it. This was submitted along with documentation relating to the website's information architecture, the goals and mission of the site, grouping and labelling of content, tree structure diagrams and design ideas for the visual appearance of the site. The third assignment related to students' evaluations of other groups' websites and finally, in the fourth assignment, the groups reform and redevelop their website based on the feedback obtained both from the tutors and from the other students on the module. The feedback they received from their work on the second assignment (the minimal content website) was therefore extremely important as this was used to guide the development of the full content website for assignment four. An example of the feedback provided for one part of the assignment is shown in table 6 below.

**Table 6:** Example of feedback comments related to marks awarded for one section of the documentation

| Mark awarded | Feedback comment |
|---|---|
| 5 | The goal and mission is superbly clear, unique and specific. Scenarios describe in depth credible potential users with credible search strategies. Analysis of competitor sites is deep. This will be of industrial standard. |
| 4 | Very good goal and mission. Good scenarios and good analyses of competitor sites. This will be high standard. |
| 3 | Unimaginative but credible goal and mission. Scenarios a bit generic but with some local colour. Analysis of competitor sites not great but generally correct. |
| 2 | Uninteresting goal and mission. No detail in the scenarios – or ones taken from a very post facto point of view. Analysis of competitor sites done only cursorily. |
| 1 | Uninteresting goal and mission. Scenario not credible or not there. Analysis of other sites purely descriptive without analysis |

 

It was interesting to compare group performance on assignments 2 and 4. It would be expected that performance on assignment 4 would be likely to improve if feedback on assignment 2 were effective. Table 7 shows a comparison of the marks for assignment 2 and 4 on this module.

**Table 7**: A comparison between the marks for group working on two related assignments, the second test (Ass.4 )took place after feedback was provided on Assignment2

| | Assignment 2 | Ass. 4 |
|---|---|---|
| Group 01 | 26 | 26.5 |
| Group 02 | 37 | 36.5 |
| Group 03 | 33 | 36 |
| Group 04 | 30 | 22 |
| Group 05 | 29 | 35 |
| Group 06 | 31 | 32.5 |
| Group 07 | 22 | 26 |
| Group 08 | 38.5 | 42 |
| Group 09 | 24 | 26 |
| Group 10 | 28.5 | 35 |
| Group 11 | 36 | 29.5 |
| Group 12 | 28.5 | 38 |
| Group 14 | 28 | 44 |
| Group 13 | 45.5 | 13 |
| Group 15 | 26.5 | 28.5 |
| Group 16 | 46 | 38.5 |
| Group 17 | 20 | 22 |
| Group 18 | 29.5 | 24.5 |
| Group 20 | 18 | 31 |
| Group 22 | 22 | 35.5 |
| Group 23 | 38.5 | 26 |
| Group 24 | 32 | 30 |
| MEAN | 30.4 | 30.8 |

Statistical analysis of the results shown in table 7 using an ANOVA showed no significant difference between the performance on the two assignments ($p<0.4$). Fourteen groups scored higher on the second assignment while eight groups scored lower. There was anecdotal evidence that some features of the assignment (the website design in particular) were better than in previous years. The results of this are hard to interpret however as multiple variables were involved. The assignment was a group assignment and group composition was a likely source of variance for example. Although it was hard to draw conclusions from this result it offers interesting possibilities for future studies.

Student comments on the nature of the feedback they received were elicited in end of module questionnaires. A selection of comments received is shown in table 8 below.

**Table 8**: Examples of student comments on the feedback received.

| Student | Comment |
|---|---|
| 01 | I got the marks back in less than a month after submitting. The feedback was very quick and email idea was good. Better than discussing the assignment in a lecture months later. |
| 02 | Comments were helpful |
| 03 | I would like all my feedback done by email. |
| 04 | The feedback document was to the point and useful to me. I could see where I went wrong and how to improve for the next assignment |
| 05 | It was clear to me how to improve next time. |
| 06 | I did not agree with everything that was written but it was very detailed and I could argue with the lecturer about my score. |
| 07 | I prefer to discuss my results person to person. I suppose this is the next best thing in such a big group. It was very fast compared to other modules and it was fair. |
| 08 | I don't read my uni email regularly so I missed my grade and the feedback till later. It would be good if it could be sent to my hotmail. |
| 09 | The feedback was simple and easy to understand. It was detailed so I could see exactly what I did right and where I went wrong |
| 10 | There was quite a lot to read but at least I have a record of it with my marks. |

A summary of the large number of comments received is shown below.

- Speed of automated feedback system was valued
- Feedback comments were in general helpful to students and of the right level of detail
- Students were able to argue about and discuss the marks awarded as the feedback was detailed enough to allow good checking (some tutors were unhappy with this).
- Feedback comments were cited as leading to improvement in work and good for revision
- Feedback provided in lectures was too general and was less valuable
- Feedback was simple and easy to understand
- Provided a permanent record which other formats did not
- Some students reported that there was lots to read; Some students evidently don't like reading
- Some preferred face to face feedback
- Email delivery was sometimes missed or spammed.
- Students didn't always agree with feedback.

 

## 9. Discussion

The research described here has presented the development, testing and evaluation of an automated feedback and marking system. The system was shown to be efficient and useful to both students and staff using the system. The evaluation of the system by staff and students suggests that the feedback quality was good and was delivered quickly and effectively using electronic mail. This was a vast improvement when compared to the manual methods used previously. The quality of feedback is vital and this was improved through modifications to the system, using the best features of automation and also by allowing tutors to add and save their own comments for later use. The use of a user-centred iterative prototyping approach involving staff and students was vital to the development of the system. It was important that the system developer was able to understand the detailed requirements and functions of the system based upon the thoughts and opinions of a range of users. In this way the system was more likely to be accepted by colleagues and external examiners and more likely to be beneficial to all. Currently the third stage prototype is being tested prior to use with learners. It is hoped that the next version will, in addition to practical and mixed practical/theory examinations, be used in a pure essay type theory test, where it is expected to be especially beneficial. It is hoped that in later stage prototypes feedback comments will be inserted directly into the documents being assessed at an appropriate place in the text. This idea is currently under consideration for prototype version four.

## References

Anderson, L. W. & Krathwohl, D. R. (Eds.) (2001) A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives. Longman, New York.

Barker, T (2009) An automated feedback system based on adaptive testing: extending the model, Proceedings of the Interactive Computer Aided Learning Conference, ICL2009, 23 - 25 September 2009, Villach/Austria

Barker, T. & Barker, J. (2002) "The evaluation of complex, intelligent, interactive, individualised human-computer interfaces: What do we mean by reliability and validity?", Proceedings of the European Learning Styles Information Network Conference, University of Ghent, June 2002.

Bloom B. S. (ed.) (1956) Taxonomy of Educational Objectives, the classification of educational goals – Handbook I: Cognitive Domain New York: McKay

Bloxham, S. and Boyd, P. (2007). Developing Effective Assessment in Higher Education. Berkshire: Open University Press.

Bridge, P. and Appleyard, R. (2005). 'System failure: A comparison of electronic and paper-based assignment submission, marking, and feedback'. British Journal of Educational Technology, 36 (4), 669-671.

Chickering A. W., and Gamson Z. F., (1987) "Seven principles for good practice in undergraduate education" *American Association of Higher Education Bulletin* vol.39 no.7

Denton, P., Madden, J., Roberts, M. and Rowe, P. (2008). 'Students' response to traditional and computer-assisted formative feedback: A comparative case study'. British Journal of Educational Technology, 39 (3), 486 - 500.

Freeman, R. & Lewis, R. (1998) Planning and Implementing Assessment, Kogan Page, London.

Hepplestone, S. and Mather, R. (2007). 'Meeting Rising Student Expectations of Online Assignment Submission and Online Feedback'. [Online]. In Proceedings of the 11th International Computer-Assisted Assessment Conference 2007, 10-11 July. Learning and Teaching Development, Loughborough University, Loughborough, UK.

Jones, D. and Behrens, S. (2003). 'Online Assignment Management: An Evolutionary Tale'. In Proceedings of the 36th Hawaii International Conference on System Sciences, 6-9 January, Big Island, Hawaii.

Lilley, M. & Barker, T. (2002) "The Development and Evaluation of a Computer-Adaptive Testing Application for English Language", Proceedings of the 6th Computer-Assisted Assessment Conference, Loughborough University, United Kingdom, pp. 169-184.

Lilley, M. & Barker, T. (2003) "Comparison between Computer-Adaptive Testing and other assessment methods: An empirical study", Proceedings of the 10th International Conference of the Association for Learning Technology (ALT-C), University of Sheffield, United Kingdom.

Lilley, M. & Barker, T. (2004). "A Computer-Adaptive Test that facilitates the modification of previously entered responses: An empirical study", Proceedings of the 2004 Intelligent Tutoring Systems Conference, Lecture Notes in Computer Science 3220, pp. 22-33.

Lilley, M., Barker, T. & Britton, C. (2004) "The development and evaluation of a software prototype for computer adaptive testing", Computers & Education Journal 43(1-2), pp. 109-123.

Lilley, M., Barker, T. & Britton, C. (2005) "The generation of automated learner feedback based on individual proficiency levels", Proceedings of the 18th International Conference on Industrial & Engineering Applications of Artificial Intelligence & Expert Systems, Lecture Notes in Artificial Intelligence 3533, pp. 842-844.

Mutch, A. (2003). 'Exploring the Practice of Feedback to Students'. Active Learning in Higher Education, 4 (24), 24-38.

Nielsen, J. and Mack, R.L. (Eds) (1994). Usability Inspection Methods, John Wiley & Sons Inc

Parkin, H. & Thorpe, L. (2009). 'Exploring student experiences of e-learning: A Phenomenographic Approach'. Paper presented at Bera Annual Conference, 2nd - 5th September 2009, University of Manchester, Manchester, UK.
Price, M. and O'Donovan, B. (2008). 'Feedback – All that effort, but what is the effect?'. Paper presented at EARLI/Northumbria Assessment Conference, 27-29 August, Seminaris Seehotel, Potsdam, Germany.
Sommerville, I. (2010). Software Engineering, Addison-Wesley, Massachusetts, ISBN-10: 0137035152
Winter, C. and Dye, V. L. (2004). 'An investigation into the reasons why students do not collect marked assignments and the accompanying feedback'. [Online]. Last accessed 12/04/2010 at: http://wlv.openrepository.com/wlv/bitstream/2436/3780/1/An%20investigation%20pgs%20133-141.pdf

 