

**MEASURING STAFF ATTITUDE TO
AN AUTOMATED FEEDBACK
SYSTEM BASED ON A COMPUTER
ADAPTIVE TEST**

Trevor Barker and Mariana Lilley

Measuring Staff Attitude to an Automated Feedback System based on a Computer Adaptive Test

Dr Trevor Barker
Mariana Lilley
Dept of Computer Science
University of Hertfordshire
Hatfield
AL10 9AB
t.1.barker@herts.ac.uk
m.lilley@herts.ac.uk

Abstract

In Higher Education today, increasing reliance is being placed upon the use of online learning and assessment systems. Often these are used to manage learning, present information and test learners in an entirely undifferentiated way, all users having exactly the same view of the system. With the development of increasingly large and complex computer applications and greater diversity in learner groups, consideration of individual differences and greater efficiency in learning and testing have become important issues in designing usable and useful applications.

Our initial findings, reported at CAA 2005, suggested that students valued this approach to providing automated feedback and considered it to be a fast, effective and reliable method. In the study presented in this paper, the attitude of staff to our automated feedback tool is presented. Three presentation sessions involving more than 80 university lecturing staff were undertaken and their views of the feedback tool were captured using video recordings. Initially a small group of computer scientists took part in a short presentation followed by a discussion where they presented their views on the CAT approach, the adaptive nature of the system and the provision of feedback. The second study involved a presentation and feedback session with more than 50 lecturers from all sectors of the university who provided their opinions of the approach in general. A short questionnaire was administered at the end of this session. The results of this, which broadly support our approach to automated feedback, are presented in this paper. A third study is reported, which involved 20 lecturers with special interests and roles in online and blended learning within the university.

Subsequent analysis of the sessions using qualitative data analysis methods showed that teachers in general were receptive to the idea of automated

feedback based on CAT. Several interesting ideas arose from the discussions, which are presented here. Computer based testing and automated feedback are becoming increasingly important in Higher Education. It is important that the views of teachers are considered when developing and implementing such systems if they are to be accepted and hence effective.

Introduction

Despite the reported benefits of the computer-aided assessment approach, high staff/student ratios often mean that tutors are unable to provide learners with feedback on assessment performance that is timely and meaningful. Freeman & Lewis (1998) amongst others have reported on the importance of feedback as a motivator for student learning. 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 who are assessed via computer-aided assessment applications.

In earlier work, we have shown that a system of automated feedback, based on student performance in a Computer Adaptive Test was useful, efficient and generally well regarded by students (Lilley and Barker 2002; 2003; 2004). Barker and colleagues (2002) noted the importance of all major stakeholders in design, implementation and evaluation of projects related to online learning. For this reason, it was important to consider also the views and attitudes of teaching staff to the provision of automated feedback based on a CAT. For this reason, three studies were undertaken to obtain detailed views and suggestions related to our automated feedback prototypes. A summary of the sessions is presented below.

Session 1 Group of 10 computer scientists, teachers, experts in software design and also interested in the provision of online educational systems. A 30 minute presentation followed by a 30 minute moderated, focussed discussion.

Session 2 Group of 50 university lecturers at university conference presentation on MLE. A 25 minute presentation followed by a 5 minute question session and a short questionnaire.

Session 3 Group of 20 university teachers interested in online and blended teaching and learning underwent a 30 minute presentation and 30 minute moderated, focussed discussion.

Each of these sessions each involved a short presentation of the automated feedback prototype, including sample output screens, examples of feedback and also research data related to student performance and attitude to the feedback provided. After each presentation, a semi-structured question and answer session was conducted, where researchers and staff could exchange ideas. Sessions were moderated by an experienced researcher and discussion topics were focussed, based upon a previously prepared script.

The sessions, however, were semi-structured, since open discussion was encouraged on any topic related to the discussion topics.

Sessions were recorded on video and later transcribed in full by the researcher and analysed, using QSR N6 software, in order collate and link together themes and ideas. Responses on the questionnaire administered in session 2 were summarised and is presented below in table 3.

Computer-Adaptive Test (CAT) Prototype Employed in this Study

The development of the CAT application that was the subject of this study has been reported by Lilley and colleagues (Lilley et al. 2004; 2005). The application comprised a graphical user interface, an adaptive algorithm based on the Three-Parameter Logistic Model from Item Response Theory (Lord, 1980; Hambleton, 1991; Wainer, 2000) and a database of questions. This contained information on each question, such as stem, options, key answer and IRT parameters. In this work, subject experts were employed for question calibration. The subject experts used Bloom's taxonomy of cognitive skills (Pritchett, 1999; Anderson & Krathwohl 2001) in order to perform the calibration. Questions were first classified according to cognitive skill being assessed. After this initial classification, questions were then ranked according to difficulty within each cognitive level. Table 1 summarises the three levels of cognitive skills covered by the question database and their difficulty range. It can be seen from Table 1 that knowledge was the lowest level of cognitive skill and application was the highest. An important assumption of our work is that each higher level cognitive skill will include all lower level skills. As an example, a question classified as application is assumed to embrace both comprehension and knowledge.

Difficulty b	Cognitive Skill	Skill Involved
$+1 \leq b \leq +3$	Application	Ability to apply taught material to novel situations
$+1 < b < -1$	Comprehension	Ability to interpret and/or translate taught material
$-1 \leq b \leq -3$	Knowledge	Ability to recall previously taught material

Table 1: Level of difficulty of questions

At the end of each assessment session, questions were re-calibrated using response data obtained by all participants who attended the session. In general terms, questions that were answered correctly by many test takers had their difficulty levels lowered and questions that were answered incorrectly by many test takers had their difficulty levels increased.

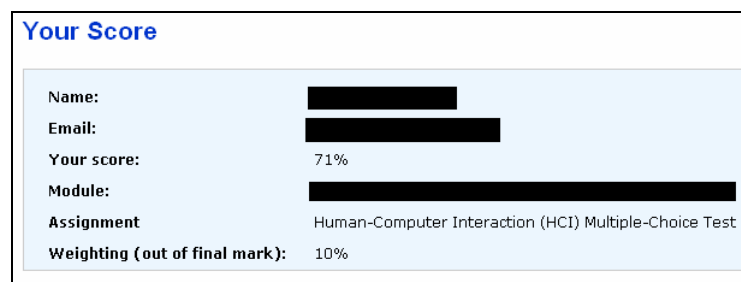
Our Approach to the Provision of Automated Feedback

It was one of our assumptions that a tutor-led feedback session would typically comprise the provision of an overall score, general comments on

proficiency level per topic and recommendations on which concepts within the subject domain should be revised. It was then planned that the feedback would be made available via a web-based application.

Overall Score

The overall score, or overall proficiency level, would be estimated by the CAT algorithm using the complete set of responses for a given test-taker and the adaptive algorithm introduced in section 2.1. Figure 1 illustrates how this information was displayed within our automated feedback prototype.



The screenshot shows a feedback interface titled "Your Score". It contains a table with the following information:

Your Score	
Name:	[Redacted]
Email:	[Redacted]
Your score:	71%
Module:	[Redacted]
Assignment	Human-Computer Interaction (HCI) Multiple-Choice Test
Weighting (out of final mark):	10%

Figure 1: Screenshot illustrating how overall score was displayed within our automated feedback prototype. The student's name and module have been omitted.

Performance Summary Per Topic

Test-takers' responses would be grouped by topic and a proficiency level calculated for each set of topic responses. Proficiency level estimates per topic would then be mapped to Bloom's taxonomy of cognitive skills. The underlying idea was to inform learners about their degree of achievement for each topic domain. Some learners reported that they would also like to compare their test performance with the performance of the whole group. This information was also made available in this section of the feedback, as illustrated in Figure 2.

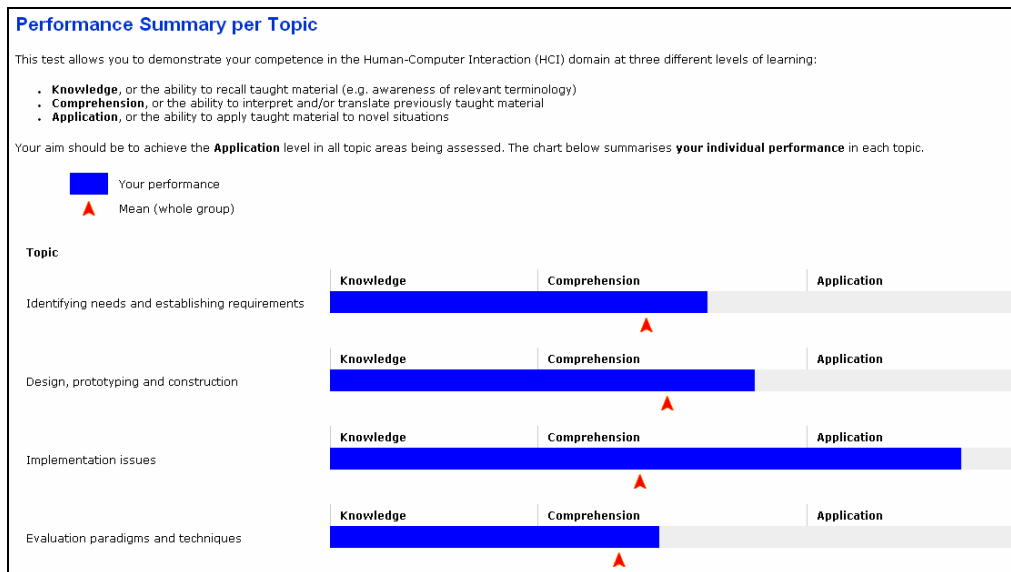


Figure 2: Screenshot of screen containing information regarding performance per topic.

Recommended Points for Revision

An important assumption of our feedback tool was that tutors providing feedback on an objective test during a face-to-face session were likely to provide students with directive feedback rather than simply indicating what the correct options for each question were. As an initial attempt to mimic some aspects of how a subject domain expert would provide learners with recommendations on how to increase their individual proficiency levels, a database of feedback sentences was designed and implemented. This database comprised statements relating to each one of the questions. For each individual student, only those questions answered incorrectly were selected. Figure 3 illustrates the approach to directive feedback employed in this study.

Recommended Points for Revision

Your personalised revision plan comprises four sections:

- [Identifying needs and establishing requirements](#)
- [Design, prototyping and construction](#)
- [Implementation issues](#)
- [Evaluation paradigms and techniques](#)

Identifying needs and establishing requirements

Did you know this?	Further action that you should do...
Use cases are usually employed for capturing the functional requirements of a system. A typical use case should convey a scenario's typical course of events.	Read Sections 7.6.2 and 7.6.3 from "Interaction Design: Beyond Human-Computer Interaction" and identify the difference between a use case and an essential use case .
Utility is a usability goal that refers to the extent to which the system provides the right kind of functionality.	Read Section 1.5.1 from "Interaction Design: Beyond Human-Computer Interaction".
'The system must support a user who is likely to be a well-trained engineer or scientist who is competent to handle technology' is an example of a user requirement .	Identify a user requirement for your Semester B [redacted] project.
A user-centered approach is characterised by "Early focus on users and tasks", "Empirical measurement" and "Iterative design".	Read Chapter 9 from "Interaction design: beyond human-computer interaction", focusing on Section 9.3 ("What is a user-centered approach?").

Back to [top](#).

Figure 3: Example of 'Recommended Points for Revision' for the topic 'Identifying needs and establishing requirements'. The module name has been omitted.

Learner Perspectives on the Usefulness of the Automated Feedback Tool

It was important to ensure that the attitude of learners to the automated feedback tool was positive. In CAA 2005, we provided a report of an evaluation of a feedback session with a group of 113 Computer Science undergraduates participated in a session of summative assessment using our CAT prototype. (Lilley and Barker, 2005). In that study, students received feedback on test performance via the automated feedback tool.

Students then completed a questionnaire in which rated a series of statements using a Likert Scale from 1 (Strongly disagree) to 5 (Strongly agree). A group of 97 students answered the questionnaire and their answers are summarised in Table 2 below.

Question	Strongly disagree 1	2	3	4	Strongly agree 5	Mean	Std Dev
Overall, the feedback tool was effective at providing helpful advice for individual development.	4	5	15	43	30	3.93	1.02
Overall, the feedback tool was effective at providing feedback on performance.	4	4	13	44	32	3.99	1.01
The "Overall Score" section was useful at providing information on how successfully I have learned.	6	9	23	31	28	3.68	1.17
The "Performance Summary per Topic" was useful at providing information on how successfully I have learned in each topic area.	6	6	19	34	32	3.82	1.15
The "Points for Revision" section was useful at providing information on how successfully I have learned.	8	9	14	35	31	3.74	1.24
Overall, I was satisfied with the degree of personalisation offered by the application.	10	7	19	35	26	3.62	1.25
The content of the feedback was appropriate for my individual performance.	6	6	20	39	26	3.75	1.11

Table 2: Learners' perceived usefulness of the feedback approach employed (N=97)

The results shown in Table 3 suggest that the automated feedback approach was favourably received by the learners who participated in the study. It was therefore important to investigate tutors' attitude towards the automated feedback approach proposed here. It was important to be sure that the approach was also acceptable to staff.

Tutors' Perspectives on the Usefulness of the Automated Feedback Tool

Questionnaires

Data obtained in the three sessions reported in section 1.1 was summarised and collated. In the second session, a short questionnaire was administered to provide information on aspects of the automated feedback approach related to formative and summative assessment, objective and essay type tests, and the speed, quality and appropriateness of the approach overall. The answers of 19 tutors who attended the presentation are summarised in Tables 3 and 4 below.

Question	Not useful 1	2	Useful 3	4	Very useful 5	Mean	Std Dev
In the context of summative assessment, the automated feedback approach that I have just seen is:	1	1	10	1	6	3.53	1.17
In the context of formative assessment, the automated feedback approach that I have just seen is:	0	0	8	3	8	4.00	0.94
In the context of objective testing (i.e. multiple-choice questions), the automated feedback approach that I have just seen is:	0	1	7	2	9	4.00	1.05
In the context of written assignments, the automated feedback approach that I have just seen is:	6	5	5	0	3	2.42	1.39

Table 3: Tutors' perceived usefulness of the feedback approach proposed in this study (N=19)

Question	Poor 1	2	Good 3	4	Very good 5	Mean	Std Dev
With regards to its <i>speed</i> , the automated feedback approach that I have just seen is:	0	0	4	3	12	4.42	0.84
With regards to its <i>quality</i> , the automated feedback approach that I have just seen is:	1	1	8	4	5	3.58	1.12
With regards to its <i>appropriateness</i> to enhance students' learning experience, the automated feedback approach that I have just seen is:	1	0	6	4	8	3.95	1.13

Table 4: Tutors' perceived speed, quality and appropriateness of the feedback approach proposed in this study (N=19)

It can be seen from tables 3 and 4 that tutors in general considered the approach to be a useful method for the provision of feedback. This is an important finding, since it will be important that tutors as well as students value the method. Table 3 shows that it is valued more highly in the context of formative, rather than summative, assessment. The use of such automated methods for written assignments was considered the least useful. It was not clear whether this was because of the difficulty of providing automated feedback for written work, or that tutors feel that providing feedback themselves was a better approach. Table 4 shows that on average tutors thought the automated approach to be fast, appropriate and of good quality, though the quality dimension achieved the lowest mean score. All in all tutors' attitude to the approach was positive, which was an important finding.

The Discussion Sessions

In all, three discussion sessions were employed in this study, based on methods described by Barker and Barker (2002). The focus of the second session was primarily to collect the questionnaire data presented in the previous section above. Accordingly there was little opportunity for discussion in this session, which contributed little to the qualitative data obtained. The bulk of the qualitative data obtained in this study therefore was collected in session 1 with a group of 10 computer scientists teachers who were also experts in software design and Session 3 involving a group of 20 experienced

university teachers who were primarily interested in online and blended teaching. In both sessions, after the presentation of our ideas and results, copies of actual feedback (made anonymous) was distributed for inspection. The discussion topics for both sessions are presented in table 5 below.

Discussion topics
What feedback methods do you use at present?
How do you assess the quality of feedback provided at present
What are the limitations and benefits of the feedback you provide currently
What is your view of the CAT approach for formative and summative assessment
What is your opinion of the CAT approach to automated feedback
What are the benefits of the approach
What are the limitations of the approach
How could the automated approach be improved
What should be the role of the tutor in the automated feedback system
What is the need for monitoring and how might this be achieved
What if any are the ethical issues in the method

Table 5: Discussion topics used in focussed sessions 1 and 3

After the presentation on the CAT automated feedback approach by one of the research team, the session moderator introduced the focussed discussion session with a short scripted introduction where the objectives of the discussion and ethical issues, such as confidentiality and the video recording were described to participants. In the first instance, the moderator started the discussion session by asking the first question in table 5, related to the type of feedback provided by tutors at present. Discussion was good in both sessions and for the most part, the moderator merely had to check that all the topics had been covered adequately, and to encourage all present to engage in the discussion where possible. When discussion moved far from the focus, or sufficient time had been spent on a thread, new topics were introduced by the moderator as unobtrusively as possible.

In the first session some discussion by the experts present was related not only to the feedback, but also to the adaptive and modelling ideas related to the software itself. This valuable information was used later, primarily to assist in the software development process in order to improve later iterations of the application. These discussions are not reported in this paper. In the following, a summary of discussions is presented under the topic headings shown in table 5.

Feedback Methods Used at Present

At present, feedback methods employed are mostly classroom and lecture theatre based sessions lasting approximately one hour, given some time after the test, ranging from six weeks to several months. Such sessions are not individual, generally each question is worked through and in some cases, general problems identified by tutors are covered in greater depth. If a question is well answered by most students, then less time is spent on this question. Problem questions are dealt with more fully by most tutors. Other

methods include providing only the questions and worked answers online (either through a web-based system, or by electronic mail). One tutor was using a spreadsheet to attempt to individualise feedback, which amounted to personally typing in comments to the answer sheet for each student. For essay type questions, feedback was usually given as comments written in pen (or sometimes electronically) onto the essay script. Sometimes feedback was provided in small group sessions where topics were discussed, rather than questions analysed in detail. One tutor reported that she used one-to-one sessions to provide feedback on rare occasions. Feedback method seemed to be related to the type of test. For objective tests, most of the methods were employed, with the obvious exception of writing directly on scripts. The purpose of feedback was very much formative, and few reported giving any feedback on summative assessments.

Quality of Feedback Provided at Present

Tutors emphasised the necessity to be able to interact directly with learners and, based upon experience, provide directed and tailored feedback. It was possible to “gauge” how a test had gone, and to provide the necessary feedback in an appropriate format. When pressed as to how this was possible, given large class sizes and the small amount of time devoted to feedback, some tutors agreed that it was not always possible. The quality of feedback provided did indeed vary according to some tutors and inexperienced colleagues might on occasions provide feedback that was variable. When asked to think about the problems of high performing and very low performing students, most tutors agreed that feedback was usually focussed at “the average” student, with an account taken of general problems that appeared in the test itself. Several tutors expressed the opinion that that the quality of our individualised automated feedback was likely to be high, citing the direct feedback on questions answered, the relationship with cognitive aspects of learning as given in the link to Bloom’s levels, and the provision of direct online links to more challenging advanced work as well as remedial work based on individual performance on the test. As the feedback was provided in a web format, links to remedial and more challenging materials were active and direct.

Limitations and Benefits of the Feedback Provided Currently

The benefits of the current system might be summarised under the possibility of direct control and monitoring of test performance and feedback. Tutors liked the ability to be able to “keep a finger on the pulse” when providing feedback. Some concern was expressed that an automated approach would lead to potential problems going un-noticed. This could not happen when tutors themselves gave feedback. Some tutors realised that un-timely feedback was far less useful than feedback given quickly. One tutor asked why we imposed a delay in giving out our automated feedback, as it could in theory be sent to students immediately after a test. The need to delay presentation of feedback due to checking and ethical reasons was less likely

to cause undue delay in the future. Most tutors agreed that the speed of the automated feedback was a major benefit.

The CAT Approach for Formative and Summative Assessment

The CAT approach was not the main focus of the discussion, as staff attitude to the CAT aspect had been the subject of earlier studies. It was important however to discuss the CAT in context of the feedback. Most staff were familiar with the CAT approach, as it has been in use in the university for several years now. Benefits of a CAT in terms of efficiency, motivation and plagiarism were already well known. Linking the feedback provided to a CAT was important for us, but not for some other tutors who could see how our automated feedback system could be linked to non-adaptive question banks, though some agreed that there would be a loss of information in such systems, related to the CAT levels in each topic area and the link to Bloom's levels. The use of CAT in summative assessment was generally less well received than for formative testing, which was in accordance with our earlier findings and the questionnaire data from session 2. It was noted by one tutor, however, that the use of a CAT for summative assessment did ensure that timely feedback would be available for all learners at the end of their course, before they had all left the university

The CAT Approach to Automated Feedback

It was realised that the use of automated feedback was an important benefit of the CAT approach. Although some tutors wanted to discuss the CAT approach in greater detail, this was resisted by the moderator and the topic of discussion gently moved. Some tutors expressed the fact that they realised that individual student profiles obtained from a CAT, containing information on performance in topic areas, as well as cognitive levels could be used in a variety of different ways. Some good ideas related to their potential use in teaching and learning were obtained from the session. Some of these are presented in the concluding section of the paper. It was noted that the use of a CAT in automated feedback involves two issues that were closely linked in our study, a CAT and automated feedback. It was our belief, expressed in the presentation, that a CAT was essential to provide individualised and rich automated feedback. It is fair to say that some tutors were not entirely convinced of this link.

Benefits of the Approach and Limitations of Automated Feedback

The most important concern expressed at the sessions related to the loss of control by tutors. Providing automated feedback was liable to remove an important "human aspect" of the teacher's role. The most important benefit cited was the speed of feedback possible with our approach. Other limitations expressed related to the use of objective testing as the only method with the approach and to issues related more to the CAT approach than the feedback itself. Other potential benefits cited included the motivational aspects of CAT

and how this might be used in order to help students do extra work, either remedially, or as extra challenges. This was seen as an important aspect by some tutors. It was emphasised in the presentation prior to discussion that the CAT level obtained represented an important boundary for an individual between what they knew and what they did not know. Providing feedback at this boundary was important and this view was expressed by some tutors present at both sessions. One teacher asked if the profiles obtained in our CAT might be useful in other subject areas. It was possible, due to the objective nature and reproducibility of CAT results that more general information related to learners might be obtained, though we could not confirm this interesting point. Efficiency of the method was also cited as a benefit. Providing feedback in traditional ways was difficult and inefficient as well as being slow. An automated system, once in operation could process test results efficiently with the minimum of human intervention. Admittedly some tutors saw this as a disadvantage, though these were in the minority at both sessions.

Suggested Improvements of the Automated Approach

There were a few suggested improvements to the system. One tutor expressed the opinion that the CAT feedback might be used as the focus either for group seminars or for small remedial classes. It would be possible to obtain useful summaries of strong and weak points in the tests in each topic area from the CAT. Such summaries might be useful to tutors in their teaching and for providing remedial materials or lectures. The speed of the CAT would be likely to provide such information quickly and certainly in time for action. Patterns of feedback might be identified in this way and the item database could be analysed to identify problem areas (and areas of strength) in all topics.

The Role of the Tutor in the Automated Feedback System

It is fair to say that a concern of some tutors was that automated feedback was another step on the road to an uncertain impersonal future. This was rarely expressed fully, though it was apparent from some questions that it was a concern. Others expressed the view that there was an opportunity in the approach to develop useful systems that would provide them with more time to develop interesting online and off-computer activities related to the outcome of tests, for example activities related to performance on tests. One teacher suggested that tests could be developed where feedback could be directly incorporated into the CAT and that this might provide a learning opportunity within a CAT. Although outside the scope of our research, this was nonetheless an interesting idea for the future. There would need to be a monitoring role as well as a development role in automated feedback systems and tutors would need to take on this aspect.

Monitoring of the Automated System: Ethical Issues

Our approach to making sure students were not disadvantaged either by our CAT approach or by the way feedback was provided in our system was explained in the introductory presentation. No tutor expressed the feeling that learners would be disadvantaged either by the CAT or by the method of providing feedback as described by us. Most stated the view that it would be important to monitor the CAT and feedback systems to ensure that they were performing properly and fairly. One tutor suggested a method of sampling, both for CAT results and feedback to ensure fairness.

In summary a complex range of issues related to the provision of automated feedback were discussed in these two focussed sessions. Additional information was obtained by means of a questionnaire, completed by attendees at a presentation related to our feedback system. In discussion sessions, tutors were able to explore a range of topics related to how feedback was provided by themselves and colleagues currently and how feedback was provided by our CAT method. In general our approach was well received and tutors were receptive to the ideas in general. They were able to see potential benefits in terms of speed and efficiency and also the ability to personalise feedback at a time when online learning is becoming increasingly important in Higher Education and staff time for providing individual feedback is decreasing. Concerns related to the provision of automated feedback were general in nature, rather than specifically directed at the system we presented. These tended to be focussed on the loss of human input into the system. There was no evidence from these sessions that feedback currently provided by tutors was of a universally high standard or that it was individualised. Rather the contrary opinion was mostly expressed.

Discussion

Substantial investments in computer technology by Higher Education institutions and high staff/student ratios have led to an increased pressure on staff and students to incorporate electronic methods of learning and teaching. This includes a growing interest in the use of computer-aided assessment, not only to make the technological investment worthwhile but also to explore the opportunities presented by the computer technology available. It is our experience that - given the great deal of computerised objective testing that currently takes place - using adaptive tests is an interesting, fair and useful way of providing such assessment (Barker & Lilley, 2003; Lilley et al., 2004). Not only is this motivating for learners, who are challenged appropriately - i.e. not discouraged by questions that are too hard, or de-motivated by questions that are too easy - but also the information that it provides can be used in interesting and useful ways. For instance, it can be used in the presentation of remedial work for students or, as in our case, for the provision of personalised feedback.

Feedback must be timely to be useful. Our experience is that when large-scale computerised objective testing is used in a formative context, results are usually returned quickly, because of automated methods of marking. Feedback, however, is often slow and delivered by the time the course has moved on and it is of less use or, in some cases, feedback is absent. This experience was largely confirmed by the results obtained in the current study. It is time consuming to produce individual feedback for hundreds of students. When feedback is provided, it is usually little more than a final score, generic worked examples and a list of questions answered correctly and incorrectly. Automated methods are therefore likely to be useful in this context, as evidenced by the tutors' attitude reported in this study. The matching of adaptive testing and automated feedback provides an opportunity to individualise feedback to a far greater extent. We argue that the automated feedback approach proposed here, which is based on adaptive testing, is appropriate for identifying learners' strengths and weaknesses for each topic area covered by the test. Automated feedback as proposed in this study is also related to Bloom's levels, thus providing meta level information for learners about the depth of their approach in each of the topic areas. This information would be difficult to obtain with standard objective testing.

Other approaches to the provision of feedback to groups of learners, such as in-class sessions where all questions from an objective test are presented by a tutor, are likely to remain as important feedback methods. Such in-class approaches offer high quality information about the test and each of the questions, often providing learners with an opportunity to work through the questions. They do not, however, address the individual needs of many of the learners. Explaining a question that is set at a difficulty level that is too low for most learners will not be of interest for the majority of the group. Similarly, it can be argued that discussing questions that only one or two learners are capable of answering will not be the most efficient way of employing tutors' and learners' time. We suggest that not only is the automated feedback based on adaptive testing a fast and appropriate method, but that it also provides information to learners that would be difficult to obtain elsewhere, given the decrease in the number of face-to-face sessions, the increase in staff/student ratios and the growing trend in the use of electronic resources for the delivery of courses, assessment, student feedback and support.

Our research has shown that learners and tutors accept and value the automated feedback approach proposed in this study. In the future we intend to apply this method more widely, for example in providing feedback for written assignments. We also intend to use the wealth of information about learners' proficiency levels provided by the adaptive testing approach to develop useful student models. Such student models will, in turn, be employed to generate profiles that could be used in a wide variety of learning contexts.

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. & 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.
- Freeman, R. & Lewis, R. (1998) *Planning and Implementing Assessment*, Kogan Page, London.
- Hambleton, R. K. (1991) *Fundamentals of Item Response Theory*, Sage Publications Inc, California.
- 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.
- Lord, F. M. (1980) *Applications of Item Response Theory to practical testing problems*. Lawrence Erlbaum Associates, New Jersey.
- Pritchett, N. (1999) "Effective question design" *In* S. Brown, P. Race & J. Bull (Eds.), *Computer-Assisted Assessment in Higher Education*, Kogan Page, London.
- Wainer, H. (2000) *Computerized Adaptive Testing (A Primer)*, Lawrence Erlbaum Associates, New Jersey.