A NEW MODE OF TEACHING FOR SELF-DIRECTED LEARNING

Eliathamby Ambikairajah[1], Julien Epps[2,1], Ming Sheng[1] and Branko Celler[1]
[1]The University of New South Wales, Australia
[2]National ICT Australia (NICTA)

ABSTRACT

Despite many advances in teaching approaches and technologies in recent years, current methods for revisiting live lecture content, which can suit the learning pace of each individual student, are not developed to their full potential. This paper describes a tool, for either capturing or pre-recording both the lecturer's video linked with the lecturer's dynamically annotated slides in digital format, using an electronic whiteboard. We report on how digitally annotated pre-recorded lectures were used in a novel mode of course delivery for self-directed study, and discuss the impact on weaker students in particular. Evaluation by both students and lecturing staff reveal this new mode of engineering teaching to be extremely promising for self-directed learning, and its demonstrated enhancement in student learning has seen its increasing use in course design and delivery.

INTRODUCTION

Collective staff experiences and regular interaction with students over many years of teaching has shown that most students suffer some degree of difficulty learning in a large class teaching environment. Challenges in this environment include the communication of difficult mathematical concepts, maintaining students' attention span, the difficulties of catering for individual student needs in a large classroom environment, different paces of student learning, lack in fluency in spoken and/or written English, and students losing continuity due to missed classes. As a result, only a small percentage of students are usually able to grasp the key concepts at the time of the live lecture delivery, while the remaining students are left to develop this critical understanding in their own time, with whatever assistance they can find and comprehend.

Educational technology has offered a range of new avenues for alleviating this problem. Brotherton and Abowd's(3) eClass automated note taking service for capture and access in the classroom has shown that multimedia capture of lecture content at the time of delivery seems to encourage student review activities that are considered helpful for performance. eClass was not reported to have a measurable impact on student performance(3).

We believe that the recording of non-persistent information such as video and dynamic annotation can be used to improve self-directed student learning or as preparatory material for more focused lecture/tutorial sessions, offering new possibilities for course delivery. Previous studies (Anderson *et al*.[2]) have shown that in practice most annotations are attentional marks, which provide critical linkage between spoken context and the slide content.

Previous experiences with an electronic whiteboard-based teaching laboratory (Ambikairajah *et al*.[1]) have shown that these teaching aids can be used very effectively to promote student-teacher interaction during tutorials, and to automatically preserve the outcomes of these tutorials in pdf format for subsequent student study. We have also introduced the technical details of the VCPlayer, a new tool for either capturing or pre-recording both the lecturer's image and annotated slides in digital format for future playback, using either a tablet PC or an electronic whiteboard (Sheng *et al*.[5]]).

In this paper, we report on how the pre-recorded lectures were used in a new mode of teaching that went beyond recording of the live lectures (Brotherton and Abowd[3]), and discuss the impact on self-directed learning. Unlike existing reports of similar tools (Ma *et al*.[4]), we have applied the VCPlayer in both undergraduate and graduate courses, and describe the evaluation of this approach herein.

COURSE DELIVERY MODES

Most universities tend to deliver courses according to one of two broad modes: (i) Full-time attendance (where students are required to attend a specified proportion of all classes) and (ii) Distance learning, or some more flexible mode that imposes minimal attendance requirements.

Within our school, background surveys have found that full-time students attending the required classes frequently do not study sufficiently out of class. Reasons for this include work commitments, fatigue in class, fast-moving large classes, and comprehension difficulties. Some students do not adapt well to this environment, and become frustrated, demotivated and prone to missing classes or dropping out.

Our student feedback has indicated that the ability to independently review live lecture content at their own pace outside of class hours would be a major benefit to them. In particular, they need to revisit lectures to solve problems arising during large classes, to alleviate language difficulties, to learn difficult concepts, and to do ongoing revision throughout the course. In turn, this facilitates their increased study commitment out of class. Remarkably, this need for self-directed study resources closely matches similar needs for distance learning and flexible learning modes.

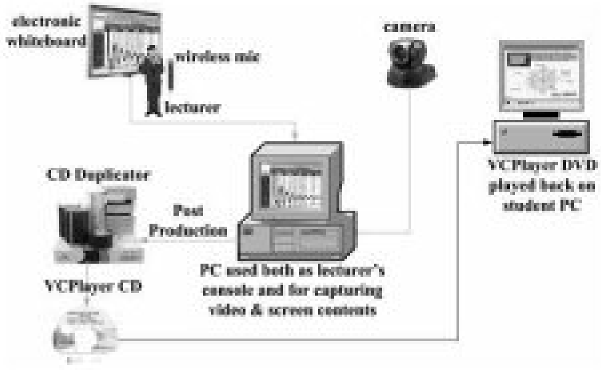


Figure 1: Live lecture capture using the VCPlayer, with electronic whiteboard

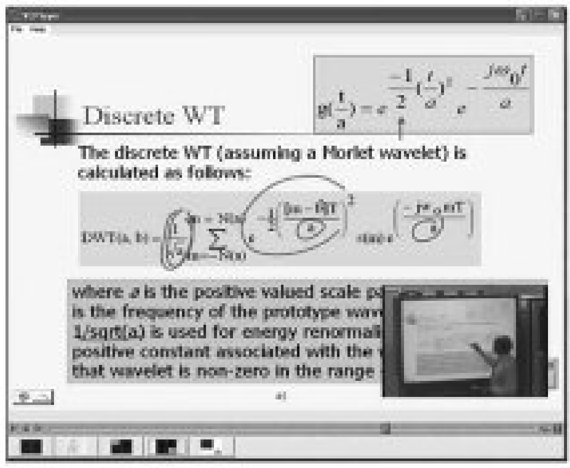


Figure 2: Synchronous lecturer video and dynamic annotation of PowerPoint™ slides on the electronic whiteboard

VCPLAYER USE IN COURSE DESIGN

VCPlayer Capture

The VCPlayer recording set-up is shown in **figure 1**, and can be performed with a PC-based video capture and an electronic whiteboard. Lectures can be recorded either in the lecturer's office (on Tablet PC), or during the live lecture (on an electronic whiteboard). Upon completion of the recording, the recorded contents are prepared for distribution on DVDs. If a high-bandwidth connection is available, the video and dynamic annotation can alternatively be streamed to students. In our trials of the VCPlayer, we have mainly made annotations on pre-prepared PowerPoint slides (as seen in **figure 2**), however blank slides can be used instead, or blank slides can be interleaved with the pre-prepared slides, as desired.

Because the VCPlayer captures whatever is being displayed on the electronic whiteboard, if needed, the lecturer can switch to other applications during the lecture. Thus, simulations and demonstrations can also be seamlessly captured, and annotated, for future playback, as seen in the example of **figure 3**.

VCPlayer Playback

The VCPlayer allows students to play back the lecture as if they were in the real classroom, with control over the view of both the lecturer video and whiteboard contents simultaneously as shown in **figures 2** *and* **3**. The lecture view window can be overlaid on top of the annotated

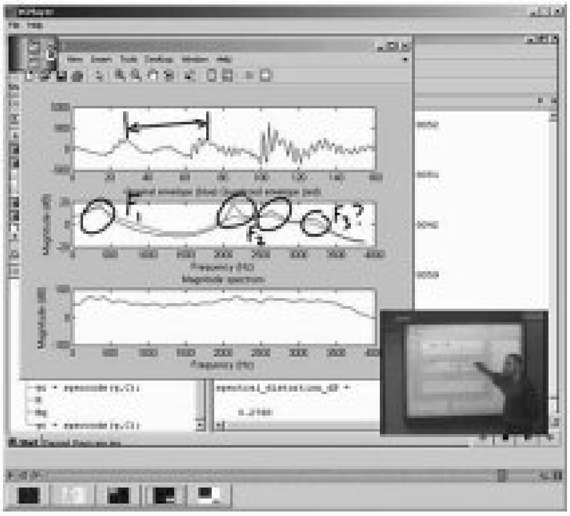


Figure 3: Playback of the lecturer's video and dynamic annotation of a MATLAB™ simulation

slides in a number of pre-determined positions using the toolbar at the bottom of the VCPlayer, or alternatively, it can be positioned in a non-overlapping position to the right of the annotated slides, to avoid obscuring the slides.

Unlike existing user interfaces (Ma *et al*.[4]), these views can be selected by the student while playback is in progress.

EVALUATION

Student Issues

As mentioned in the course delivery modes, background surveys on issues affecting student learning were conducted in 2003, 2004 and 2005. The background surveys were completed by around 300 students in total (all undergraduates), and revealed that 75% of students do less than 10 hours per week of study after class (during semester, excluding assignments and laboratory preparation and write-up), and 36% do less than 5 hours per week. The university expects students to do at least one hour of study for each contact hour of lecture or tutorial, i.e. around 3 hours lecture and 0.5 hours tutorial per week per course. 14 hours per week (minimum) study for a typical load of four courses is considerably more than the 5-10 hours we found most students were committing, and is a major motivating factor behind recent efforts to encourage self-directed study.

Objectives and Procedure

The objective of the evaluation presented in this paper was to gauge the impact of the VCPlayer-based course design on students' learning, the influence on their learning behaviour, and the extent of their acceptance of it as a mode of delivery.

In the third-year undergraduate class (Signal Processing and Transform Methods), 10 lectures (of 12 total) were given in the traditional live format, and two lectures (two hours each) were distributed as pre-recorded CDs with a discussion class replacing the live lecture. The discussion classes were highly informal, promoting intense student-lecturer interaction, either on an individual basis or a group basis. The reasons for not giving more lectures using the proposed mode of delivery were that the students should compare the traditional lecture format to this style of teaching, and that this was the students' first exposure to this style of teaching.

In the postgraduate class (Speech and Audio Processing), seven lectures (of 12 total) were given in the traditional live format, five lectures (two to three hours each) were distributed as a pre-recorded CD with a discussion class (1 hour) replacing the live lecture, and one lecture was given in the traditional format but recorded using the VCPlayer during class and distributed for revision purposes after the class.

The main evaluation was conducted using a survey at the conclusion of the courses, which was completed by 71 undergraduate students (of 129 enrolled) and 24 postgraduate students (of 24 enrolled) who were present in class at that time. The survey comprised 16 questions in multiple-choice format (mainly a 5-point Likert scale) and an open-format question for general comments. Such end-of-term methods of evaluation may not produce generalisable results, and are no substitute for ongoing evaluation and student interaction, however during the session, considerable amounts of informal feedback were also solicited from students.

Results – Learning Experience

Among the students surveyed, nearly all used the pre-recorded CD at least once, 79% of whom used the CD several times in the course of their learning. The students found that the pre-recorded lectures on the CD (plus a follow-up discussion class) were an acceptable alternative to a live lecture (73% agreed). The majority of students felt that having the pre-recorded CDs improved their learning experience (85% agreed).

Results – Offline Lecture Review and Live Follow-up as a New Teaching Mode

Responding to the use of pre-recorded lectures with a follow-up discussion class, students were generally very positive. Specific results include:

- The majority of both groups of students felt that it was more convenient for them to learn via a pre-recorded CD than attend a live lecture.
- 75% of students felt that they would be more likely to watch 100% of the lectures if they were on a pre-recorded CD than attend live lectures with no CD.
- 43% of undergraduate students found that the pre-recorded CDs were a more efficient way to learn than attending live lectures, whereas 54% of postgraduate students were unsure of whether there was a difference in the learning efficiency of both teaching methods.

Results – Level of Understanding: Live vs Pre-Recorded

Students had more opportunities to ask questions in classes having watched the pre-recorded lectures than in live lectures without the CD as shown in **figure 4**.

55% of the students liked the opportunity to review the pre-recorded lectures on CD because live lectures generally produce a more challenging learning environment. Challenges in the live learning environment include language difficulties, the fact that students have only just been exposed to the course material in the class, and the confidence required to ask questions in a formal, large class environment. Of the student population, 95% felt that the use of the pre-recorded CD allowed them to learn at their own pace, as seen in **figure 5**.

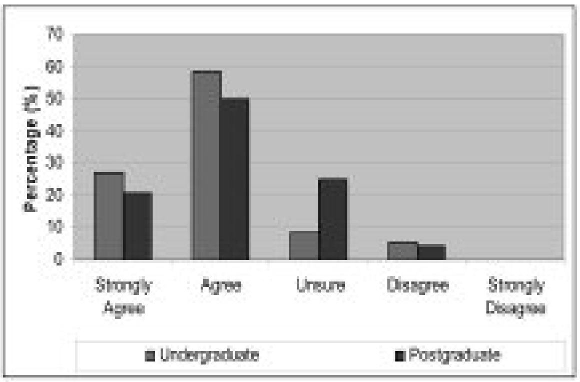


Figure 4: I was able to achieve my desired level of understanding for topics using pre-recorded CDs only

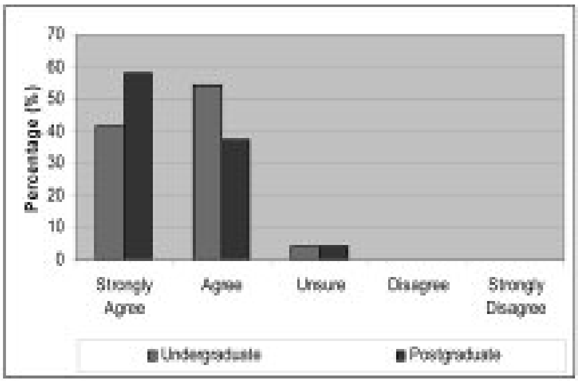


Figure 5: I liked the ability to review CD-based lectures at my own pace to improve my understanding

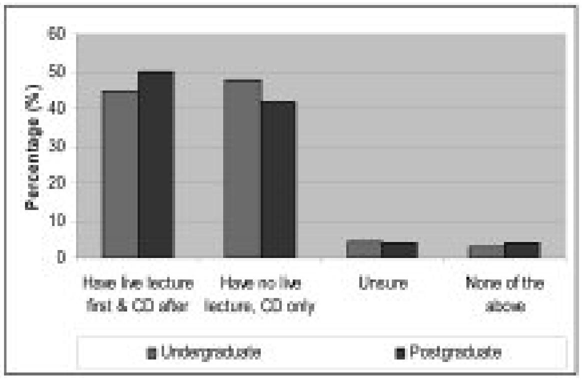


Figure 6: CD-based course delivery preferences

Results – Pre-Recorded vs. Live Lecture Capture

When it came to a preference between having a live lecture first and receiving a pre-recorded CD afterwards or receiving a pre-recorded CD first, having some discussion in class later, but having no live lecture, there were notable differences between undergraduate and postgraduate students, as seen in **figure 6**.

Student and Staff Feedback

One of the key results from the student survey was that 80% of students recommended the use of pre-recorded CDs (followed by class discussions) for other courses within the school, although undergraduates were marginally more enthusiastic for this than postgraduates.

In order to optimize the VCPlayer, staff found that well-developed PowerPoint slides were a distinct advantage, although in most cases staff had such slides prepared already. Since the introduction of the VCPlayer in 2005, 20% (6 out of 30) of the staff in the school are using it in their lectures. The ease of use of the VCPlayer has seen initial interest in adoption by staff from other parts of the faculty. An unexpected benefit was that staff occasionally wanted to use the review functionality of the VCPlayer in order to check that specific details had been fully explained. Finally, once recorded, a lecture can potentially be re-used as a resource from one year to the next.

Educational Outcomes

Students' questions on the subject matter for the week were generally observed to be sharper and reflected a deeper understanding in discussion classes (after watching the CD).

In terms of academic performance at the completion of the course, an improved pass rate was observed for the same level of paper over a 3-year period for an undergraduate cohort, as seen in **table 1**. This concurs with the informal observation that fewer weaker students appeared to be struggling with the course material.

Year	No. of students	Failure rate (%)
2003	121	19.0
2004	118	13.6
2005	129	10.1

Table 1. Student performance using live lectures only (2003), live lectures and some tutorial solutions using AVI files (Ambikairajah *et al.*[1]) (2004), and live lectures together with pre-recorded CDs (student preferred) (2005)

CONCLUSION

This paper has presented an application of the VCPlayer, a novel tool for the synchronous capture of live lecturer video and electronic annotation and playback. An evaluation of student learning in undergraduate and postgraduate courses designed to exploit the functionality of the VCPlayer has revealed that students and staff responded very positively to its ease of use, review capability and new educational possibilities. Initial indications of the migration from traditional lecture formats in 2002-3 to the format enabled by the VCPlayer in 2004-5 show that the student cohort's understanding of key signal processing concepts is improving, evidenced by a gradual increase in student grades. Future research and careful monitoring will be required to rigorously substantiate this trend. This approach also lends itself to applications in distance learning.

ACKNOWLEDGMENT

This project was funded by the Australian Government Department of Science and Training (DEST) Capital Development Pool Programme 2004-2005 (Education Delivery).

REFERENCES

1. **Ambikairajah, E., Epps, J., Sheng, M, Celler, B. and Chen, P.**, 2005, '*Experiences with an electronic whiteboard teaching laboratory and tablet PC-based lecture presentations*', in Proc. IEEE ICASSP, vol. V, pp. 565-568.
2. **Anderson, R. J., Hoyer, C., Wolfman, S. A. and Anderson, R.**, 2004, '*A study of*

digital ink in lecture presentation', in Proc. Conf. on Human Factors in Comp. (CHI), pp. 567-574.

3. **Brotherton, J. A. and Abowd, G. D.**, 2004, '*Lessons learned from eClass: Assessing automated capture and access in the classroom*', ACM Trans. Computer-Human Interaction, vol. 11, no. 2, pp. 121-155
4. **Ma, M., Schillings, V. and Meinel, C.**, 2003, '*T-Cube: A multimedia authoring system for eLearning*', in Proc. AACE E-Learn, pp. 2289-2296.
5. **Sheng M., Celler B., Ambikairajah E. and Epps, J.**, 2005, '*Development of a virtual classroom player for self-directed learning*', in Proc. Int. Conf. on Multimedia and Inf. & Comm. Tech. in Education (Spain), pp. 718-721.

EVALUATION OF THE RELIABILITY OF A NEW SCHEME FOR CONVERTING MCQ TEST SCORES TO PERCENTAGE MARKS

Yuyuan Zhao
The University of Liverpool, United Kingdom

ABSTRACT

Multiple choice question (MCQ) tests are a widely used assessment methodology. In summative assessments, MCQ raw scores must be converted into more meaningful marks or grades. This paper introduces a recently developed conversion scheme and evaluates the applications of the scheme to an engineering computing module over two academic years. The assessment of the module consisted of 5 MCQ tests. The conversion scheme was applied to each test and then the marks of the 5 tests were averaged to give the final module marks. There was a reasonable correlation between the students' marks of this module and their overall average marks of all the other modules taught in the same semester. The changes in subject difficulties and improvement in teaching method were evidenced in the changes in the average marks of the individual tests. The MCQ assessment and the application of the conversion scheme are proved to be successful.

INTRODUCTION

Multiple choice question (MCQ) tests are a widely used assessment methodology. They are objective, easy to mark and quick to obtain results. Properly designed MCQ tests are an efficient tool for assessing a large number of candidates, especially in knowledge or fact based subjects.

Designing MCQ tests and interpreting the test results are not as straightforward as setting and marking conventional examination papers. A conventional examination paper consists of a relatively small number of questions, which are explicitly or inexplicitly divided further into smaller elements. Each element is assigned a fixed mark. The mark that a student obtains in the examination is simply the sum of the marks of the correctly answered elements. The mark is normally expressed in percentage and is considered to be a true measure of the students' performance. Any MCQ test, however, invites some guesswork and therefore involves some uncertainty in the test results. This uncertainty issue must be addressed in summative assessments using MCQ tests.

The reliability of a MCQ test depends on the suitability of the subject content for MCQs, the format of the questions, and the method adopted for converting the marks. For a MCQ test to serve as a reliable and effective assessment method, three criteria must be met.

Criterion 1: each question is properly designed as a suitable measure of a learning outcome. This criterion is not always easy to be met. The main reason is that MCQ assessments are not suitable for some subjects. Engineering education involves acquiring knowledge, describing processes and phenomena, solving problems, developing experimental methodologies, creating designs, deriving mathematical formulas, analysing data, performing calculations and developing communications skills etc. It is difficult, if not impossible, to formulate MCQs for some of the above elements. For example, it is extremely difficult to split a design project into small and independent assessable problems. Even if MCQ assessments are suitable for the subject, badly devised MCQs can impair the effectiveness of the assessment. A good MCQ should have equally feasible choices of answers and the correct answer cannot be identified by a layman. The quality of the design of MCQs is largely dependent upon the skills of the question setter, as subject relevance is concerned.

Criterion 2: the format of the MCQ test is designed to ensure that the fluctuation in the test scores as a result of random guesses is minimised. Zhao(1) analysed the effect of the format of a MCQ test on the part played by guesswork on the test scores. The probabilistic analysis confirms that the optimum number of choices of answers for each MCQ question is four(1). In a test composed of questions with two- or three-choices, there is a high chance of obtaining a pass score by pure guesswork and the scores often fall within a narrow range. It is often difficult to differentiate the students' performance. Introducing five or more choices of answers does not offer significant benefits in reducing the effect of guessing. From a subject point of view, questions with five or more choices of answers are also difficult to construct. The probabilistic analysis shows that for any given type of MCQs, the number of questions in a test has a determining effect on the reliability of test scores and thus test marks(1). For a four-choice question test, the minimum number of questions needed to reduce the probability of obtaining a mark above 40 by pure guesswork to below 5%, 1% or 0.01% is 8, 18 or 48, respectively. While the more questions the better, 20 four-choice questions are sufficient for ensuring a high level of reliability. MCQ tests containing 10 or more four-choice questions are often good enough.

Criterion 3: raw test scores are converted to marks that are a true measure of the students' performance. McLachlan and Whiten(2) differentiated marks, scores and grades and pointed out that raw scores of MCQ tests should not be used directly. Instead, scaling should be applied before the marks of the individual assessment units are aggregated. In well established tests involving a large number of participants, such as TOEFL, complex scaling schemes are often used. These schemes are developed on the basis of extensive research on the statistics of the past tests, as demonstrated in Wainer and Wang(3). While these schemes are good for gauging the relative competence of the candidates, the converted test scores are not compatible with the percentage marks normally used in engineering assessments. Recently, Zhao(4) developed a scheme for converting MCQ raw scores to conventional percentage marks based on probability theory. The conversion scheme is independent of class size and historical data. It removes the guesswork element so that the converted marks become a true measure of the students' knowledge and competence. The converted marks are compatible with the standard marking scheme. MCQ tests can therefore be used standalone or as units of an assessment including conventional assessment units.

The author has used a series of computer based MCQ tests as the sole summative assessment method for a 7.5 credit module, *Introduction to Computing*, for the past two academic years. This module is compulsory for all first year students in the Department of Engineering, the University of Liverpool. The aim of the module is to equip the students with key computing skills for engineering applications. Traditionally, the module had been assessed by a series of coursework plus computer based tests conducted on a one-assessor-to-one-student basis. Because of the large number of skills to be assessed and the large number of students taking the module, the assessment was extremely time-consuming. It also had some shortcomings, such as difficulty in detecting copying and cheating, and long delay in giving feedback to students. The MCQ tests were introduced with an aim to achieve efficient and responsive assessments without sacrificing effectiveness and reliability.

This paper demonstrates the procedure of applying the conversion scheme developed by Zhao (4) and evaluates the outcomes of the applications of the scheme to the *Introduction to Computing module*.

IMPLEMENTATION OF CONVERSION SCHEME

The simplest representation of the conversion scheme is in the form of conversion tables. **Table 1** is for converting raw scores to standard percentage marks for MCQ tests with questions of two, three, four or five choices of answers (4). A raw score represents the percentage points a student scores in a MCQ test before any conversion is applied, and is simply termed 'score' in this paper. A standard percentage mark represents the percentage points a student deserves in a MCQ test, and is simply termed

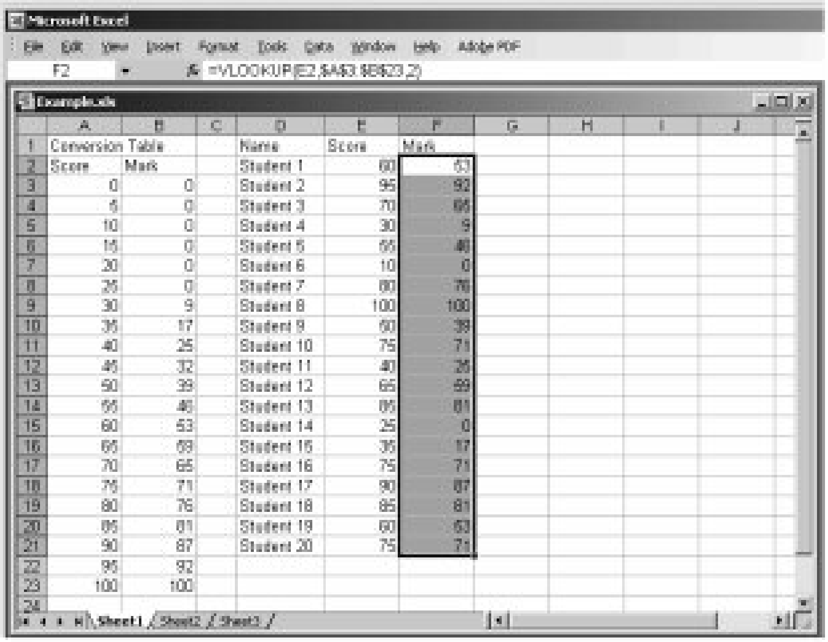
Score	Mark [2]	[3]	[4]	[5]	Score	Mark [2]	[3]	[4]	[5]	Score	Mark [2]	[3]	[4]	[5]
≤20	0	0	0	0	47	0	22	35	43	74	38	60	69	75
21	0	0	0	2	48	0	23	36	44	75	40	61	71	76
22	0	0	0	4	49	0	25	38	46	76	41	62	72	77
23	0	0	0	5	50	0	26	39	47	77	43	64	73	78
24	0	0	0	7	51	2	28	41	48	78	45	65	74	79
25	0	0	0	9	52	3	29	42	49	79	47	66	75	80
26	0	0	2	11	53	5	31	43	51	80	48	68	76	81
27	0	0	3	12	54	6	32	45	52	81	50	69	77	82
28	0	0	5	14	55	8	33	46	53	82	52	70	78	83
29	0	0	7	16	56	9	35	47	55	83	54	72	79	84
30	0	0	9	17	57	11	36	49	56	84	56	73	80	84
31	0	0	10	19	58	12	38	50	57	85	58	74	81	85
32	0	0	12	20	59	14	39	51	58	86	60	76	83	86
33	0	0	14	22	60	15	41	53	59	87	62	77	84	87
34	0	1	15	24	61	17	42	54	61	88	64	79	85	88
35	0	3	17	25	62	18	43	55	62	89	66	80	86	89
36	0	4	18	27	63	20	45	56	63	90	68	81	87	90
37	0	6	20	28	64	22	46	58	64	91	70	83	88	91
38	0	8	21	30	65	23	48	59	65	92	72	84	89	92
39	0	9	23	31	66	25	49	60	66	93	74	86	90	92
40	0	11	25	33	67	26	50	61	67	94	77	87	91	93
41	0	12	26	34	68	28	52	62	69	95	79	89	92	94
42	0	14	28	36	69	30	53	64	70	96	82	90	93	95
43	0	16	29	37	70	31	54	65	71	97	85	92	95	96
44	0	17	31	39	71	33	56	66	72	98	88	94	96	97
45	0	19	32	40	72	35	57	67	73	99	92	96	97	98
46	0	20	34	41	73	36	58	68	74	100	100	100	100	100

Table 1: Conversion table for MCQ tests with questions of two, three, four or five choices of answers, corresponding to columns indicated by [2], [3], [4] and [5]

'mark' in this paper. Take four-choice question tests as an example. A student obtaining a score equal to or below 25 is awarded a zero mark. A score of 60 corresponds to a mark of 53.

The conversions of scores of an MCQ test can be carried out by using a spreadsheet programme such as Microsoft Excel. For demonstration purposes, let us assume that a class of 20 students have taken a MCQ test composed of 20 four-choice questions, each of which is worth 5 points. The possible scores the students can obtain vary from 0 to 100 in steps of 5. **Figure 1** demonstrates how a list of scores of the class is converted into a list of marks by using Excel. Firstly, we enter the conversion table in columns A and B from row 3 to row 23, with the scores in column A and their corresponding marks in column B. Secondly, we enter the student names in column D and their scores in column E, starting from row 2. Thirdly, we use the VLOOKUP function to perform the conversion. This is realised by entering the following formula in cell F2: =VLOOKUP (E2,$A$3:$B$23,2). The formula searches the value of E2 in column A and returns the corresponding value in the same row from column B. Finally, we select cell F2 and pull down the filler handle so that all the scores in column E are converted and the marks are displayed in column F.

In practice, the arrangement of the data in Excel can be improved to give better presentation and clarity. For example, the conversion table may be entered on a separate worksheet. The same conversion table can be applied to the scores of a number of tests entered in different columns.



	A	B	C	D	E	F
1	Conversion Table			Name	Score	Mark
2	Score	Mark		Student 1	60	53
3	0	0		Student 2	95	92
4	5	0		Student 3	70	65
5	10	0		Student 4	30	9
6	15	0		Student 5	55	46
7	20	0		Student 6	10	0
8	25	0		Student 7	80	76
9	30	9		Student 8	100	100
10	35	17		Student 9	50	39
11	40	25		Student 10	75	71
12	45	32		Student 11	40	25
13	50	39		Student 12	65	59
14	55	46		Student 13	85	81
15	60	53		Student 14	25	0
16	65	59		Student 15	35	17
17	70	65		Student 16	75	71
18	75	71		Student 17	90	87
19	80	76		Student 18	85	81
20	85	81		Student 19	60	53
21	90	87		Student 20	75	71
22	95	92				
23	100	100				

Figure 1: A screen snapshot showing the conversions of scores to marks by using Excel

EVALUATION OF ASSESSMENT

The *Introduction to Computing* module was composed of 5 units: MS Word, Excel Basics, Excel Optimisation, MATLAB Basics and MATLAB Programming. Each unit was assessed by a one-hour MCQ test with four-choice questions. Each test required the students to independently perform a series of operations using the software package being tested. In the test, the students were allowed to consult training materials. The module mark for each student was obtained by converting the scores of the 5 tests into standard percentage marks followed by averaging these marks.

In the current academic year (2005/06), the marks of three modules taken concurrently by all the first year students in the first semester have been made available. The average score before conversion and average mark after conversion of the *Computing* module have been compared with the average mark of the other two modules for 231 students. **Figure 2** shows the relationship between the scores of *Computing* and the average marks of the other two modules. **Figure 3** shows the relationship between the marks of *Computing* and the average marks of the other two modules. While the class average of the average marks of the other two modules is 54.3, the class averages of the scores and marks of *Computing* are 73.3 and 67.7, respectively. Fitting the data to a straight line also shows that the scores and marks of *Computing* are on average 28% and 19% higher than the average marks of the other two modules. As the other two modules are examination based, it is understandable that the marks of *Computing* can be considerably higher than the marks of these two modules. However, using the scores of MCQ tests directly would result in unacceptably high marks. The effect is especially pronounced for weak students. As a consequence, using scores directly would result in an increase in pass rate from 93.5% to 97.8%. The data reinforces the point that MCQ test scores must be converted to standard marks.

Figure 4 shows the histogram of the differences between the marks of *Computing* and the average marks of the other two modules. 36% of the students have a difference within 10 marks, 63% within 20 marks and 85% within 30 marks. It should be pointed out that the correlation between *Computing* and the other modules in academic year 2005/06 is not as strong as that in academic year 2004/05 (1). The higher degree of scatter is mainly because of the fewer modules available for comparison in 2005/06.

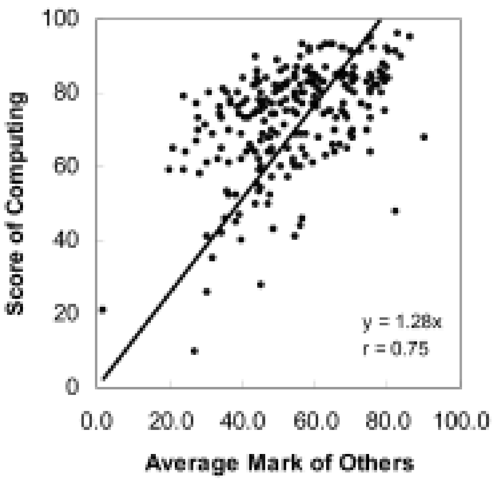


Figure 2: Relationship between the scores of *Computing* and the average marks of the other two modules

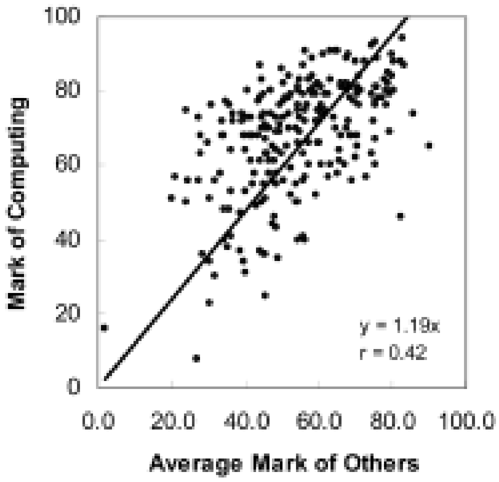


Figure 3: Relationship between the marks of *Computing* and the average marks of the other two modules

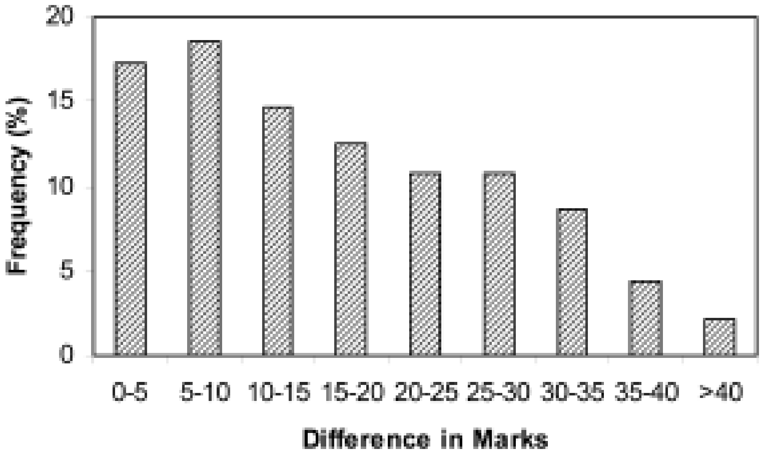


Figure 4: Histogram of the differences between the marks of *Computing* and the average marks of the other two modules taken in the same semester

Table 2 compares the average marks of the 5 individual MCQ tests in *Computing* between academic years 2004/05 and 2005/06. In 2004/05, there were big variations in the average marks among the 5 tests, which correctly reflected the relative difficulties of the 5 topics. In 2005/06, lectures and demonstrations were introduced to improve the students' learning of the difficult topics. The difficulties of the test questions were also adjusted. As a consequence, the overall module average was increased and the variations in the average marks among the 5 tests were significantly reduced.

SUMMARY

Computer based MCQ tests were used as a summative assessment method for an engineering module over two academic years. The conversion scheme developed by Zhao(4) was applied to each of the 5 MCQ tests. The final module marks were obtained by averaging the marks of the 5 tests. There was a reasonable correlation between the students' marks of this module and their overall average marks of all the other concurrent modules. The changes in difficulties of topics and the improvements in teaching and assessment methods were

Test	2004/05	2005/06
Word	82.1	79.3
Excel I	72.1	66.6
Excel II	60.8	73.5
MATLAB I	62.1	73.0
MATLAB II	49.9	54.5

Table 2 Average marks of individual MCQ tests in 2004/05 and 2005/06

evidenced in the changes in the average marks of the individual tests. The MCQ assessment and the conversion scheme were proved to be as reliable as the traditional assessment methods.

REFERENCES

1. **Zhao, Y**., 2006, *Int. J. Eng. Edu.*, 22, (in press)
2. **McLachlan, J. C. and Whiten, S. C.**, 2000, *Med. Edu.*, 34, 788-797
3. **Wainer, H. and Wang, X.**, 2001, *TOEFL Technical Report TR-16: Using a new statistical model for testlets to score TOEFL*, Education Testing Services, Princeton, USA
4. **Zhao, Y**., 2005, *Int. J. Eng. Edu.*, 21, 1189-1194.

IMPROVING MARKING EFFICIENCY AND ENGINEERING UNDERGRADUATE FEEDBACK USING INNOVATIVE SOFTWARE TOOLS

Barry Beggs, Elaine Smith and Allan Pellow
Glasgow Caledonian University, United Kingdom

ABSTRACT

The study reported in this paper describes the development and application of a software tool used to enhance the learning, teaching and assessment of engineering undergraduates. The tool was designed to support academic staff by enabling them to provide undergraduates with marks and detailed electronic feedback soon after they submitted their assessments. The feedback comments were aligned to the engineering subject benchmark statements. By supporting the creation of marking templates and combining a data base of re-useable feedback comments with the ability to add free text, an innovative solution emerged where academic freedom was not compromised by the resulting improvements in marking efficiency. As part of the study, assessor-generated computer produced feedback was investigated as an alternative to hand written feedback comments exposing some of its advantages and disadvantages. Both detailed analysis and anecdotal evidence from lecturers and students are presented. It was concluded that although there is potential to improve the consistency, quality, quantity and efficiency of providing feedback to students by using computer generated information facilitated by bespoke software tools, there are also limitations that accompany such an automated process. The tool continues to develop as part of an online submission system that includes plagiarism checking and statistical analysis of student feedback.

INTRODUCTION

We are reminded that assessment should be valid, reliable, transparent and authentic, Race(1). It should also be designed to be 'fit for the purpose' for which it was meant and, stripped down to its simplest form, this means application of academic judgement and provision of feedback. It is clearly necessary in assessing students to discover their knowledge and understanding of the content and concepts in a syllabus and the attainment level for each individual of learning outcomes. It is equally important to facilitate students to be reflective and to give them the opportunity to improve and to develop. Anecdotally, students complain that they do not get enough feedback or that the feedback comes when it is too late. The importance of feedback to students should not be underestimated and research undertaken at Nottingham Trent University showed that there is a need for feedback if the objective is that students are to learn effectively(2). They concluded that feedback is necessary for students to allow them to feel valued and 'listened to', to take ownership of their own learning and to develop reflective thinking.

It is interesting to note that although the size and demographic of the student population has changed over recent years, there is evidence from the QAA enhancement themes that methods of assessment are largely unchanged; final examination with essay or technical reports(3). Some attempt has been made to maintain effectiveness and introduce efficiency using Computer Assisted Assessment (CAA). For those involved in the design of CAA systems, Bloom's suggested six levels of educational objectives have provided a reasonable starting point(4). Questions may be categorised and mapped to learning and educational objectives and banks of questions can be built up as electronic resources. One such bank, which provides a considerable electronic resource, was developed by the Electrical and Electronic Engineering Assessment Network (E3AN) project(5). There are also a number of centres of expertise in online assessment including The Scottish Centre for Research into On-Line Learning and Assessment (SCROLLA)(6), where the objectives have been to explore the

Background	5 marks
Simulation	10 marks
Hardware	5 marks
Interpretation of Results	5 marks

Table 1: Lecturer's Marking Categories

applicability and suitability of different forms of technology to enhance the assessment process. It must be agreed that using technology, significant progress has been made in the design, presentation and marking of coursework and examinations. Unfortunately, less has been done to facilitate the other important element of assessment which is the provision and delivery of feedback to students.

The software tool described in this study was developed specifically to provide undergraduates with detailed electronic feedback soon after they submitted their assessment. The methodology was to analyse the conventional hand written feedback of a lecturer, develop and then use a software tool to re-mark the same batch of student work. The software tool would then be used by other lecturers under different circumstances and then redeveloped based on their evaluation.

ANALYSIS OF HAND WRITTEN MARKING

Before using the software tool to mark coursework, an analysis was carried out on conventional hand written feedback that had been done in a previous academic year. A feedback form was partially completed by the student on submission and attached to their technical report; thereafter the lecturer added hand written comments and provided a mark. The analysis was carried out on the content and frequency of comments on a sample of 41 submissions. This analysis identified marking categories and feedback comments included in an electronic marking tool called ELF (Electronic Lecturers' Feedback). ELF was then used to re-mark the same batch of student work.

The lecturer had initially created a marking template with four marking categories and allocated part marks to each category as shown in **table 1**. In addition to a break down of marks, the lecturer provided two or three sentences of hand written feedback comments. Examination of the forms revealed that there were 27 different comments in the reports. The most often used comment was re-written 9 times, however 13 comments were only used once. A histogram showing the frequencies of use of the 27 comments is provided in **figure 1**.

DEVELOPMENT OF THE SOFTWARE TOOL

The early version of the software tool, ELF, was specified with the intention of streamlining the provision of feedback to students. Developed in Visual Basic, it was a stand-alone executable programme and the students' copy was designed to be similar to the existing coursework feedback form.

A marking template was prepared to include relevant information, such as module and programme codes, and a selection of marking categories from a list of 25 that had been collated as a result of the earlier analysis, research and personal experience, **table 2**. Part marks could be specified by the lecturer for each marking category used and the total mark was calculated and displayed automatically.

Each of the 25 marking categories had an associated group of feedback comments that could be chosen by the lecturer as they were marking. A sample of the feedback comments available for the category 'Introduction to Report' is given in **table 3**.

REMARKING WITH THE SOFTWARE TOOL

Since no free text capability was available in the early version of ELF, marking categories were chosen by the lecturer and the 27 feedback comments were added to the appropriate group. This process had to be completed before marking could begin and was timed and took 20 minutes. Once the feedback comments were added, they could be chosen during the marking and it was concluded that it was possible to offer a rapid customisation of ELF.

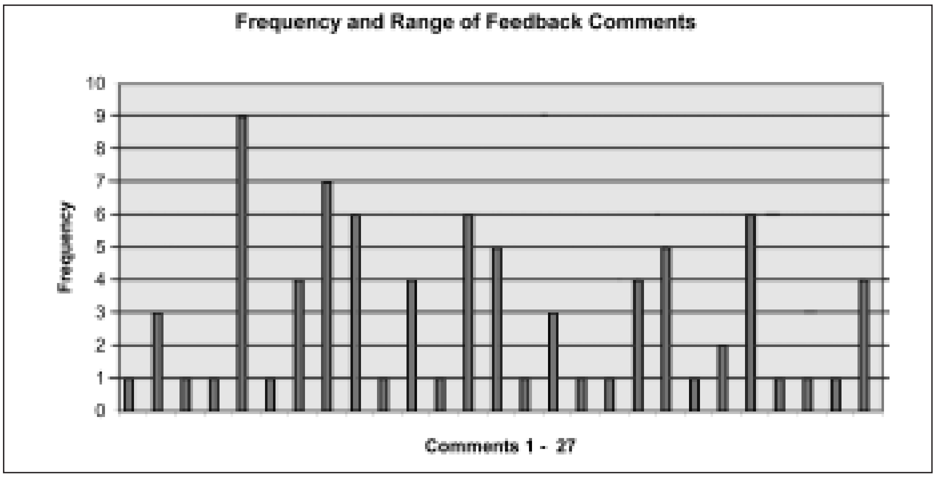


Figure 1: Feedback comments

ID	Title	ID	Title
1	Introduction to Report	14	Evaluation of Results
2	Introduction to Essay	15	Attainment of Learning Outcomes
3	Motivation, Initiative and Adaptability	16	Attainment of Benchmark Statements
4	Self Organisation and Time Management	17	Attainment of Tasks
5	Knowledge and Interpretation	18	Process and Keywords
6	Conceptual Design	19	Oral Communication Skills
7	Creativity and Innovation	20	Written Communication Skills
8	Depth of Understanding	21	Use of Diagrams/Graphs
9	Synthesis of Knowledge	22	Main Body
10	Group Work Skills	23	Word Limit, Punctuation and Spelling
11	Technical Competence	24	Referencing and Plagiarism
12	Application of Theory	25	Conclusions
13	Discussion of Limitations		

Table 2: Initial marking categories

Your introduction gave essential definitions of the aims and objectives of the report. Your introduction re specified the aims and objectives of the report. Your introduction interpreted the aims and objectives of the report, making them manageable. Your introduction sign posted the sequence of achieving the aims and objectives of the report. Your introduction only summarised the aims and objectives of the report. Your introduction to the report could have been expanded. Your introduction to the report was confusing and disorganised.

Table 3: Feedback comments for marking category 1

When asked to comment on the time taken to customise and mark the reports with the software tool compared with hand written marking, the lecturer stated 'although the general perception was that it takes a long time to write out comments by hand, in practice it was difficult to estimate just how long this takes since comments are often written during the reading and evaluation of the report'. He believed that the smooth continuous process of hand writing helped to maintain continuity and the time taken to write the comments provided him with time to reflect on, and assess the overall standard of the work.

EVALUATION BY CASE STUDIES

ELF in its early version was used by one lecturer who was keen to evaluate it for long term use with future cohorts. His initial attempts proved less successful than anticipated due to difficulties in customisation of marking categories and feedback comments. Although it was expected that the whole process would initially take longer than usual, the process became far more of a major undertaking and the decision was made to abandon its use and provide feedback forms manually. These forms were returned to students during the first available class. Discussions with the students revealed that they were keen to receive feedback like this and that it was uncommon to get feedback from lecturers. The lecturer was encouraged by the students' response and later used ELF for another class with more success. His comments included; 'It would be fair to say that the first experience with ELF was not successful and the process took longer than expected. The second attempt proved more successful'.

Another lecturer, intending to have a totally electronic process for coursework submission and provision of feedback, used ELF for a small group of students. The students had to produce a technical report as a Word or RTF document and submit it to the lecturer's e-mail address with the module code and their name in the 'subject' field. After the hand-in deadline passed he sent out e-mail 'receipts' to the students. The lecturer then used ELF to mark the work by opening two windows on his computer screen – one with ELF and the other with the coursework, allowing him to read through the work and mark it as he went along. He commented; 'I'm not sure that I saved much time with this first serious use of ELF, but I'm fairly confident that the process is workable and will become more efficient as I use it a few more times. I plan to handle another coursework electronically using ELF in the same way'.

FURTHER DEVELOPMENT AND USE

After its use during one academic year, ELF was upgraded to become a more flexible tool which combined the ability to select frequently used feedback comments from a data base with the ability to edit those comments and add free text. This further evolution of ELF also resulted in it being merged into an online software tool called KELPIE. KELPIE was developed at Glasgow Caledonian University to support the management of student absence. The ELF element of KELPIE continues to be used by a number of lecturers to provide feedback on a wide range of subject areas and assessments.

One interesting growth of the ELF element of KELPIE has been its use in marking final year engineering student project reports. The overall assessment consists of three elements; an interim report, a poster presentation and a final written report. Two programme streams deliver similar modules involving individual student projects to BEng and BSc students for various engineering subjects. The current trial involves the use of ELF for marking the interim report submitted at the start of Semester B. A total of 80 students are supervised on an individual basis by approximately 27 academic staff. The interim report is a summative assessment element of the project and it was agreed that a formalised marking template and an improved feedback process was necessary to enable students to receive the type of standardised, prompt, helpful and timely feedback as recommended by the QAA Enhancement Themes(3). **Table 4** shows the agreed marking categories that were used to generate a marking template.

The feedback comments associated with each marking category were loosely written to align

ID	Title	Mark	ID	Title	Mark
1	Presentation of Report	20	4	Self Organisation and Time Management	20
2	Critical Evaluation of Progress	20	5	Motivation, Initiative and Adaptability	20
3	Identification of Problem Areas and Solutions	20	6	Supervisor Comments	0

Table 4: Project interim report marking categories

ID	Text	Order
1	It was clear that your report is very well presented with no errors.	Excellent
2	Your report was well presented with few errors.	Good
3	Your report was reasonably presented but had some errors. You can learn how to improve your presentation skills by consulting the project handbook or at: http://www.learningservices.gcal.ac.uk	Threshold
4	Unfortunately your report was not very well presented. You should make an attempt to improve this before you write your final year project report. Advice is available from the project handbook or at http://www.learningservices.gcal.ac.uk	Below Threshold

Table 5: Feedback comments on presentation

with the QAA engineering benchmark statements. Those shown in **table 5** relate to Presentation of Report and are intended to align as indicated.

The marking template was used to guide the project supervisor to provide part marks and feedback comments for the five marking categories and to use the free text capability to add supervisor comments. This allowed the supervisor to enter part marks and standardised feedback comments to maintain consistency but to retain the individual academic freedom to include advice and judgemental statements.

OVERALL CONCLUSIONS AND OBSERVATIONS

Hand written comments created by a lecturer provide accurate feedback that can be finely tuned to convey appropriate detail and level of importance as well as a personal approach often appreciated by students. It is unlikely that any software tool could ever provide anything other than a restricted range of comments. Although this is a limitation for an individual assessor, restricting the range of comments can be a way of achieving consistency in marking and feedback provision. This may be particularly valid for project work or dissertations or when marking templates are created for large student groups assessed by different lecturers.

The early use of ELF showed that using software tools for marking was time consuming, particularly when used for the first time. The need to create comments before they could be used exasperated early users but the response from the students suggested that they were very happy to receive the amount and detail of feedback that was provided. This inspired a further development of the software tool to allow the lecturer to edit feedback comments and to add free text. Expansion of the idea of electronic feedback into the receipt of coursework in electronic form opened up additional benefits that had not been anticipated in the early development stages of the software tool.

Candid observation confirms that using bespoke software tools to support lecturers in the marking and provision of feedback to students is not without its problems. Perhaps the real value of computer-generated feedback will only be achieved when it is set within a larger context exploring where it can provide added value. This could be realised by tying it into other electronic facilities such as automatically screening for plagiarism or using it to generate statistics. There is clearly an opportunity to continue to seek ways of aligning student feedback to benchmark statements.

REFERENCES

1. **Race, P.**, 2001, *Assessment: A Guide to Students*, LTSN Generic Centre, York
2. **http://www.cap.ntu.ac.uk/se/why.html** (accessed online 20/02/05)
3. **Reflections on Enhancement**, Volume 1, QAA, Scotland
4. **Bloom, B. S.**, (ed) 1964, *Taxonomy of Educational Objectives*, The Classification of Educational Goals, V2, Longman, London
5. **http://www.e3an.ac.uk/** (accessed online 20/02/05)
6. **http://www.scrolla.ac.uk/** (accessed online 20/02/05)