FULLER, RICHER FEEDBACK, MORE EASILY DELIVERED, USING TABLET PCS

Paul McLaughlin, Wesley Kerr, Karen Howie

Fuller, Richer Feedback, More Easily Delivered, using Tablet PCs

Wesley Kerr, Karen Howie and Paul McLaughlin
(paul.mclaughlin@ed.ac.uk).
Institute of Molecular and Structural Biology,
Michael Swann Building,Mayfield Rd,
Edinburgh, EH9 3JR. 0131-6507060

Abstract

We have developed a method to use tablet PCs to enable markers more efficiently to give written feedback on students' work. Comments may either be made in handwriting, or may be typed, or may be presented in type following handwriting recognition. Additionally, any comments so made can be stored and reused, allowing for editing. Importantly, feedback can be made richer by including forward links for students to follow up on common mistakes that they have made so that their engagement with feedback is more constructive. Such feedback would otherwise be very tedious to provide if marking on paper was used exclusively.

We have run this system successfully for two years to mark essays in a large class of 450+ students, using twenty markers. This volume of work was efficiently handled and involved no paper. Checking of marks and assuring consistent standards was much more easily done than with paper.

We consulted students and markers. Students take the system in their stride. They are well able to provide essays, with diagrams and figures. Markers fell into a number of groups. We have learned that there are a variety of marking styles and developed the software to accommodate these. The only software required is Microsoft Word and Excel.

The problem addressed

Good quality feedback is the most single powerful influence on student achievement in higher education (Hattie, 1987). But a number of surveys with students shows that satisfaction with feedback on assessment is the least of all areas considered. (Hounsell et al, 2005, Krause et al, 2005, Surridge, 2006, Hounsel et al, 2007). Several reasons contribute. Too long a gap between submission and feedback is detrimental and a source of dissatisfaction (Gibbs &Simpson, 2004) Crook et al (2005) have evidence from focus groups that students sometimes simply cannot read a marker's handwriting. They also found that students considered tick sheets and/or boxes in which the marker makes comments to be too formulaic.

As Crook et al point out, many of these problems stem from a rise in student numbers that are not matched by a proportionate rise in staff, such that marking becomes a burden not a teaching opportunity. Marking and returning work for large classes indeed takes much time and resources, both for academic and administrative staff. Traditionally this is done on paper, which has the drawback that handing it back to students causes problems. Hounsell (1987) shows that many students don't pick it up. It sometimes goes missing (perhaps maliciously). If it is collected marked work often goes into a drawer, or is otherwise misplaced, such that the student can't find the work when preparing for a subsequent essay. Submission of word-processed work onto a Virtual Learning environment (VLE) might seem to solve many of these problems, but it creates others. More discursive work, such as the traditional essay, is frustrating to read on-line as most screens are not large enough to display an A4 page at sufficient resolution. This entails much tiring scrolling. Even if the marker has a large enough screen it is rarely portable, and so doesn't fit in with the way most markers work with paper copies. Marking on line also means that feedback must be typed. This becomes very tedious and especially frustrating in the sort of exercise where the marker often has to make much the same comment on many students' essays, or a make a comment that is only slightly edited from student to student

A proposed solution

This paper shows an attempted solution to some of these problems using two features of tablet PCs. These machines look like ordinary laptop computers, except that the screen can be swivelled to lie flat such that the keyboard is hidden underneath. Then the screen can display in portrait mode, as opposed to the usual landscape view, such that the screen is similar in size to a sheet of A4 paper. Indeed, a page of a Microsoft Word document can be displayed a page at a time at sufficient resolution to be easily read. and figures are similarly as readable as on paper. The second unique feature used is that the tablet is supplied with a stylus that can be used to write on the screen. The stylus effectively annotates the displayed document in "virtual" ink, again at sufficient resolution that it appears to be similar to writing on paper. Importantly, under Windows XP Tablet operating system there is handwriting recognition such that the writing input by the stylus may be converted into text. Using Microsoft Word Macros, we developed these features into a system to mark submitted work.

Implementation – first iteration

The software has been developed and used to mark essays in a large first year biology class in the University of Edinburgh. The class has roughly 480 students, each of whom submits an essay on a topic associated with evolution. The task is designed to promote students to find material to support their arguments, to help them to appreciate the differences between what is expected at school and at university, and to challenge the misconceptions that many still have about evolution (a pastiche would be

“Giraffes grew long necks to be able to eat leaves on tall trees”, but often the argument appears in essays in a more subtle form).

After completing their essays, students load them, containing their associated diagrams and figures, as Microsoft Word files onto a VLE (WebCT). These are bundled into zip files and downloaded onto Tablet PC machines, which are distributed to each marker. A “Shortcut” icon on the desktop takes the marker to an Excel file, which control the work flow. A macro button populates the file with a list of students.

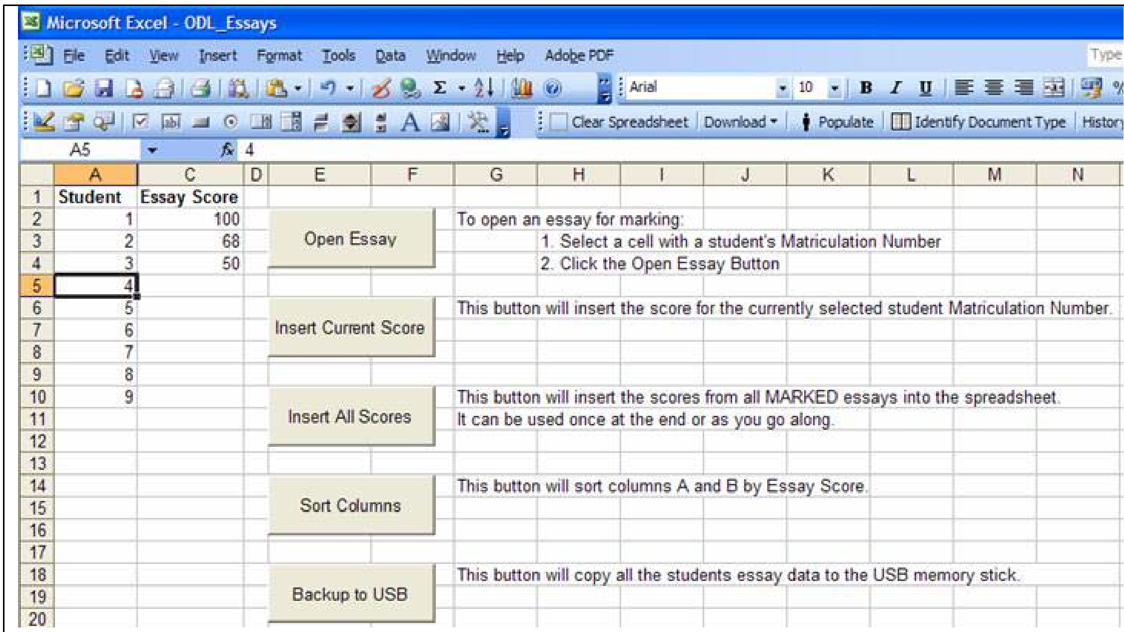

Figure 1. The Excel spread sheet used by markers to work through their assigned marking.

Each student may be selected (Figure 1), and on pressing another macro button, the student’s file is opened by calling an instance of Microsoft Word. For the purposes of marking, a particular template has been developed with a number of new toolbars and macros to facilitate marking (Figure 2). The most significant is a “Enter Comments” toolbar, which allows the insertion of comments, their storage and/or reuse.

Figure 2. The "Comments" dialogue box used to inserted preassembled feedback.

Previously stored comments may be sorted by frequency of previous use, or may be searched by keyword. They may be edited again to provide a more appropriate comment for a particular student. These comments are inserted as in "balloons" in the right hand margin, as they use the same "Comments" tool provided in Word.

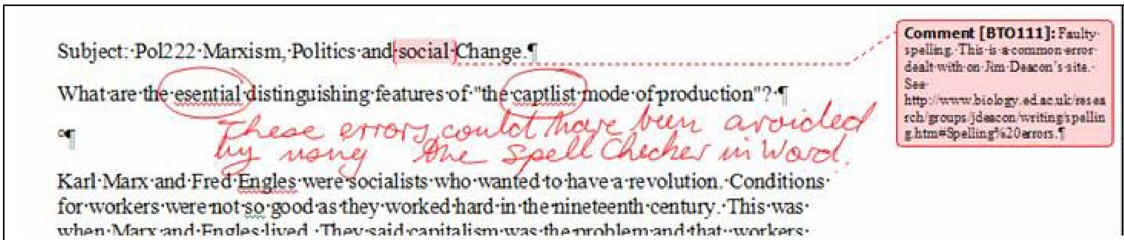

Figure 3. Examples of typed and handwritten feedback. The example essay, full of deliberate mistakes, comes from NorthernTerritories Universities' "Study Skills On Line" Site (James et al, 2002)

The pre-stored comments are an opportunity to provide students with links to remedial action. Like many HE institutions, the School of Biological Sciences in the University of Edinburgh has a website on generic skills, such as the elements of writing essays, a site on statistics and a site on spelling and grammar. Some of the preassembled comments have links to these inserted. The idea here was to both publicise these sites for students, and also hope that a student who needed correction on any one point might be drawn to other content via the link.

When the student's work is opened a mark sheet is automatically appended with the marker's name written in and with a space for a mark or grade. When the marker finishes marking, the file is closed. The marker then returns to the Excel spreadsheet and may push a macro button that causes each marked file to be visited, any mark to be read, and then the mark to be inserted into the Excel file. Thus the marker learns how far through the list they have progressed. By this method they are less unlikely to miss an essay out, or misplace a mark than if they had to transcribe marks themselves.

When all markers have returned their machines, their original Excel files are ignored. The directories containing the marked essays are bundled into a new directory structure. Then a similar Excel file reads all students' marked files, the marks and markers' name into itself. Again, marks are read from the files that will be returned to students. This is an important point because it deals with a situation whereby a marker might update the mark on the essay, but forget to update it on the Excel file. Ultimately the marker's own sheet is only of relevance to the marker to track where they are in the list of students to be marked: the student's marked essay is the "golden copy" always.

A master Excel spreadsheet controls all subsequent administration. It is used to look at markers' averages and is used to prepare new bundles of marking to be reassessed by more experienced markers for those markers who have egregious averages. It is also used to make new bundles to be reassessed for those students who are borderline fails, or so that the work of students who were not known to be special needs at the time of marking can be revisited. The spreadsheet is also populated with submission dates so that lateness penalties can be flagged. Those students who attract penalties for plagiarism are also flagged.

Finally, when all work that should be reassessed has been returned, the marked work is moved to a secure website that is protected by the university's authentication system. A dynamic link is released to the students that parses the directory name from their User Identification on WebCT and allows the student to access their own marked work and no-one else's. Such systems are not essential. It would be relatively simple to modify the Excel code to send the marked work by e-mail.

Implementation – second iteration

After the first year roughly 1300 unique comments that had been created by markers were available. The subject of the essay changes every year, so it was desirable to have only generic comments (468) to be used in subsequent years. It was decided to divide these into folders, to reduce the length of each list. The folder names were: General Comments; Spelling and Grammar; Introduction, References; Graphs, which included comments about diagrams, figures and graphs, and Conclusions. After removal of almost identical comments, but still allowing different ways of saying the same thing, the total number of comments was roughly 160.

For the next year handwriting onto the essay was introduced in addition to typed on handwriting recognition. In the first year, all comments were made by typing or by using handwriting recognition. From feedback from markers it was clear that some markers found this frustrating. Therefore in the second year we introduced markers to using the stylus directly onto the submitted work. Additionally, in the first year some markers found scrolling with the stylus to be frustrating as the mapping from the stylus to the vertical scroll bar at the edge of the screen was not accurate enough – it was also frustrating for left-handed people who found stretching across their own field of view to be annoying. Thus we decided to buy a mouse with a scroll-wheel for each machine, and this seemed to eliminate these complaints.

Evaluation

In the first year we paid particular attention to the markers' experience. It was they on whom the greatest burden of dealing with new and unfamiliar software fell, while for students little new demands were made. The most significant difference between this first iteration and the second is that we informed the markers that they could handwrite on the essays, whereas in the first year we led them to believe that handwriting recognition was the only way they could make comments. This was a deliberate deceit because we wanted to capture all comments in machine-readable form so that we could build up a database of comments to form a new list of generic comments for the next year. A second reason was that using handwriting recognition is initially slower than handwriting. We wanted to see if markers would progress in their skills at handwriting recognition and we felt that if an easier option was given, many would not persevere.

It became clear that a significant number found handwriting recognition very frustrating and that marking roughly twenty essays each was not long enough to make sufficient progress. In the second year, 4/10 markers who replied made comments exclusively in their own handwriting, while the other 6 used handwriting recognition or a mixture. Where specific markers chose to identify themselves, there was no clear correlation between either computing confidence or age with use of handwriting exclusively. Some unconfident users, who in the first year complained bitterly about handwriting recognition, used it exclusively in the second without protest. On the other hand, some younger tutors preferred handwriting exclusively.

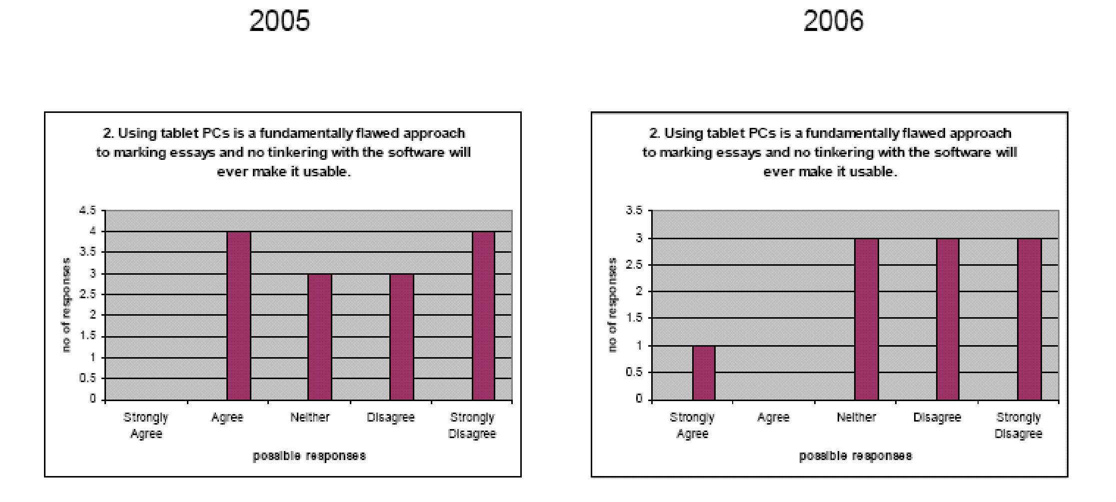

Figure 4. Histograms showing the numbers of markers agreeing with the statement that the system was fundamentally and irredeemably flawed, comparing when use of handwriting recognition was the only method to mark (2005) and after (2006) when handwriting on the work was introduced.

After handwriting was given as an option, there was a significant reduction in those that agreed with the null hypothesis (Figure 4), namely that the exercise was a "fundamentally flawed approach to marking essays and that no amount of tinkering with the software will ever make it useable"

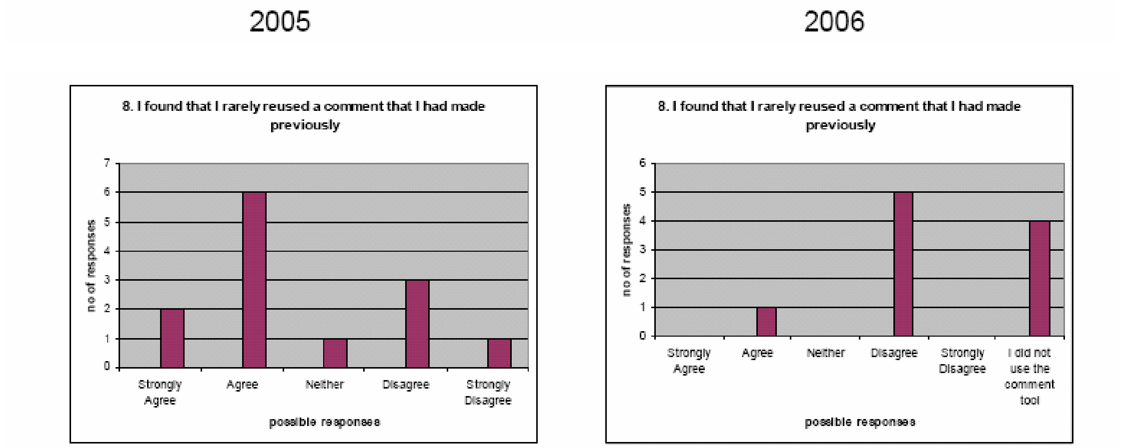

Figure 5. A Histogram of markers rejecting the null hypothesis that they rarely reused comments.

There was also an improvement in those that reused comments (Figure 5). This might have been because in the second year there was a richer bank of generic comments to use, derived from real comments made by the markers themselves in the previous year. In the first year preloaded comments were sparser and were invented abstractly rather than based on marking real essays. It might equally well have been that they did not use the comments tool bar.

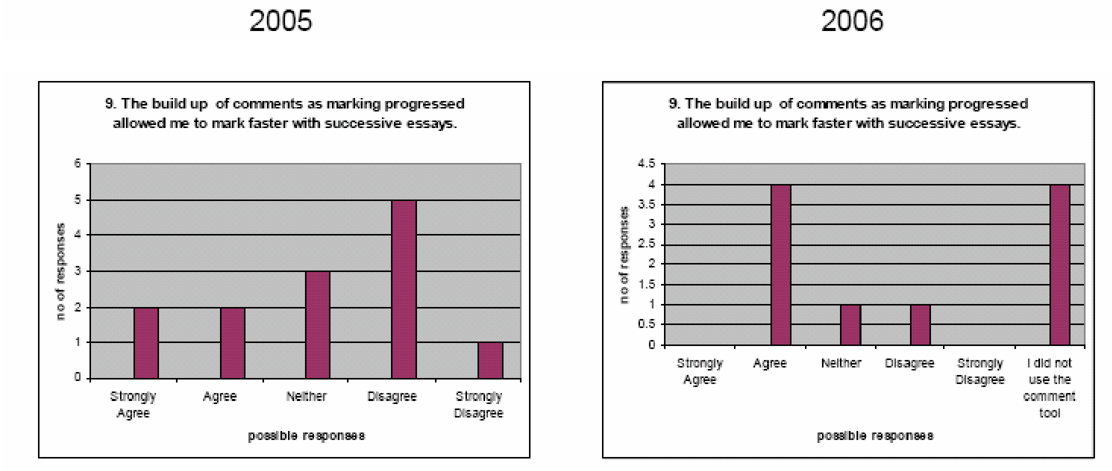

Figure 6. Histograms showing markers agreeing that the build of comments during marking helped them to mark more quickly.

It was also clear that there was an improved perception that the build up of comments during marking was improving the speed of responding as marking progressed (Figure 6).

In the second year, we conducted a survey of student reactions to the essay feedback. We had no baseline to compare improvement against. Not for the first time, a technological development leads to wider reflection on what was normal practice before the innovation. More students had their expectations of the amount and quality of feedback met or exceeded than were disappointed.

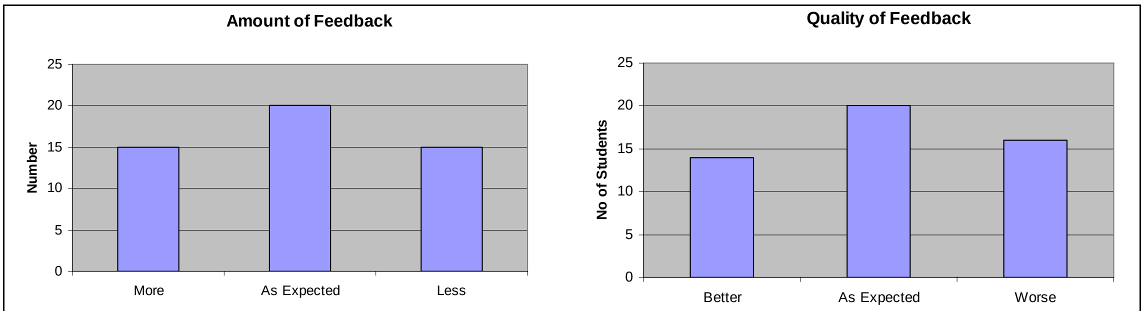

Figure 7. Histogram of students' reflections on the amount and quality of feedback they obtained.

However, it should be stated again that these were first year students and this essay was the first many had done at University. Thus these data do not disentangle the effect of the technology from orthogonal factors, such as their expectations from school or disappointment in their grades.

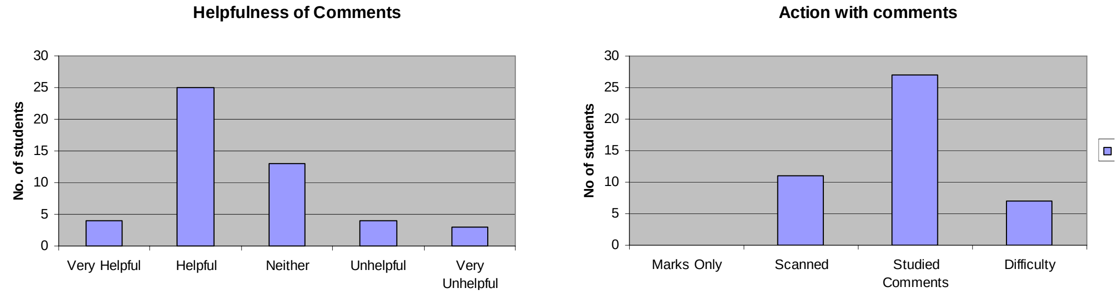

Figure 8. Histogram of students' evaluation of the helpfulness of feedback (left), and the action they took with it. "Marks only" means that they read the mark but didn't look at the feedback; "Scanned" is when they just skim-read the feedback; " Studied comments" means that they worked through them carefully; "Difficulty" means that they had difficulty making sense of the comments.

The students' perceptions of the helpfulness of the comments were more positive than negative (Figure 8). The diligence with which the students studied the comments seems gratifyingly high, although we only have their own word for this. Some had difficulty with understanding comments, either because they found the content too sparse or the meaning too elliptical, or because the simply could not read handwritten comments. There did not seem to be much gross difference between the distributions for students whose marked work had typed only comments, handwritten only comments and those that had a mixture. However when the replies were broken down in this way, there were not enough replies to be sure of seeing subtle but significant differences.

We also asked for more discursive feedback from students.

Is there any comment you would like to make on the feedback of your ODL essay?

Individual comments showed that those that were unhappy were usually dissatisfied for a reason unconnected with the technology. For example, there was a perception that it was realistic to attain a mark of 100% and that a marker's role was merely to take marks away, rather than award them. By the same token, we cannot ascribe any positive comments to being solely due to the technology.

- *"Feed back was more detailed than expected"*

- *“Very good feedback, very helpful as a different style is required for a scientific essay than I was used to, so good instructions on this have been given”*

- *“Excellent feed back, very helpful would like [another electronically marked exercise] also to have been marked in this way”.*

There was a clear theme that handwritten comments were often hard to decipher.

- *“It would be good to make all the comments typed, or at least make the markers write in capitals, as I could not decypher the handwriting of my marker.”*

- *“..couldnt read some of the comments made. handwriting was too difficult to read.”*

- *.” I assume the marker used a piece of equipment that allowed handwritting to be shown in a word document. i have re-read the feedback and still cannot make out some of the comments. The feedback given maybe very constructive but i have no way of knowing.”*

We have no data to say that these comments might also have been made on a paper version. The quality of handwriting on the electronic version does not seem to be lower resolution. In the next implementation, however, we will make the default for handwriting a “Biro” rather than a broader-nibbed “felt-tipped” pen.

There were surprisingly few comments on technical difficulties. A worrying case was:

- *“The feedback should also be available to download on macs”*

The College of Science and Engineering at the University of Edinburgh is predominately Microsoft based and Mac computers are in a minority, such that we would have had difficulty accessing a Mac computer to test. An obvious solution will be to save the marked essay as a PDF file.

In summary, it was clear that markers were now much more enthusiastic and positive about using PC tablets now that the software catered for a variety of marking styles. Just as students have a variety of learning styles, markers also have preferred ways of working. Any successful marking engine must cater for these because it is crucial that all take part. The software is still new to many of them and it will be interesting to see how the use evolves and if handwritten comments decline in favour of typewritten ones, particularly as experience grows.

Conclusions and Perspectives

The system eases administration of large quantities of feedback and should narrow the time gap between submission and receiving feedback. More, but not all, feedback is typed. As the majority of feedback becomes typed issues with legibility should reduce. Also, typed feedback can contain hyperlinks to remedial material on grammar, spelling, dealing with data and the like. A disadvantage of the system, however, is that it relies on tablet PC machines. It would be much more widely applicable if it worked on any Windows machine. Certainly, larger LCD screens that allow an A4 page to be read in one screenful are becoming cheaper, so the need for a portrait screen is lessened. Cheap graphics tablets are available but we have found none that captures handwriting at sufficient resolution. If this problem will resolve itself in the future, the present system could be run with any machine that has Microsoft Word, provided that handwriting recognition software could also be used.

In implementing the system it has been crucial to bring staff along. To this end it is important to develop software that is as flexible and as non-prescriptive as possible. Few academics like to be told that they can no longer do something that they are used to doing. We relied on the goodwill of our colleagues to take the system up. Thus we made it clear that no preloaded comment need be used and if it were used that it should be editable by the marker. Similarly the marker was free to make his/her own remarks and to store them for future use. For the same reasons we eventually “allowed” handwritten comments as well as handwriting-recognition and typing. Interestingly some markers who were vehemently against handwriting recognition in the first year, when it was the only mode of entry, used it in the second in the knowledge that they could have handwritten comments if they had wanted to.

As feedback from students showed, we are not the first to propose a computing solution to find that there are deeper pedagogic reasons for the problem we hope to solve but find that at best we can only mildly alleviate. But this method of marking assists reflection on practice because, by its nature, it accrues large amounts of data that would have been tedious to collect if we had used paper only. Thus we have in machine readable form data to sift for examples of good practice. This might have more weight with markers in the knowledge that it comes from their peers. Future areas to look at are whether novice markers are helped by having a database of remarks that more experienced markers have used, and whether this makes marking more consistent. It would also be interesting to research if more experienced markers feel more of a social pressure to give fuller feedback now that the remarks they make are stored and are seen by their peers, not only by the student being marked. Clearly it is impossible that technology in marking has a neutral effect, but not all changes are necessarily worse.

References

Crook C, Gross H and Dymott,T (2006) Assessment relationships in higher education: the tension of process and practice *British Educational Research Journal* **32**, 95-114

Gibbs G & Simpson C (2004) Conditions Under Which Assessment Supports Students' Learning. *Learning and Teaching in Higher Education* **1**, 1-31

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

Hounsell, D. (1987) Essay writing and the quality of feedback, in J.T.E. Richardson,M.W. Eysenck & D. Warren-Piper (eds) *Student Learning: research in education and cognitive psychology*, Milton Keynes: Open University Press and Society for Research into Higher Education

Hounsell D et al (2005) *Enhancing Teaching-Learning Environments in Undergraduate Courses: Final Report to the Economic and Social Research Council on TLRP project L139251099*, Universities of Edinburgh, Durham and Coventry: Enhancing Teaching-Learning Environments in UndergraduateCourses Project www.tla.ed.ac.uk/etl/docs/ETLfinalreport.pdf (accessed 11/05/2007)

Hounsel D et al (2007) Integrative Assessment. Managing assessment practices and procedures. Guide no 4 http://www.enhancementthemes.ac.uk/documents/IntegrativeAssessment/IAManaging.pdf [accessed 11/05/2007]

James,R McInnis C and Devlin M (2002) Sample Essay "LearnLine", Charles Darwin University, Northern Territories, Australia. www.learnline.cdu.edu.au/studyskills/as/Sample_essay.pdf [Accessed 10/05/2007]

Krause K et al (2005) *The First Year Experience in Australian Universities: Findings from a Decade of National Studies*, Melbourne: University of Melbourne, Centre for the Study of Higher Education http://www.dest.gov.au/sectors/higher_education/publications_resources/profiles/first_year_experience.htm (accessed 11/05/2007)

Surridge P (2006) *The National Student Survey 2005: Findings,* Bristol: Higher Education Funding Council for England www.hefce.ac.uk/pubs/rdreports/2006/rd22_06/ (accessed 11/05/2007)