Using digital pens to expedite the marking procedure

Konstantinos Banitsas*, Georgios Eliopoulos and Linda Murray

Department of E&CE (H222), School of Engineering and Design, Brunel University, Uxbridge, Middlesex UB8 3PH, UK.

E-mail: konstantinos.banitsas@brunel.ac.uk, geliop@gmail.com, Linda.Murray@brunel.ac.uk.

*Corresponding author

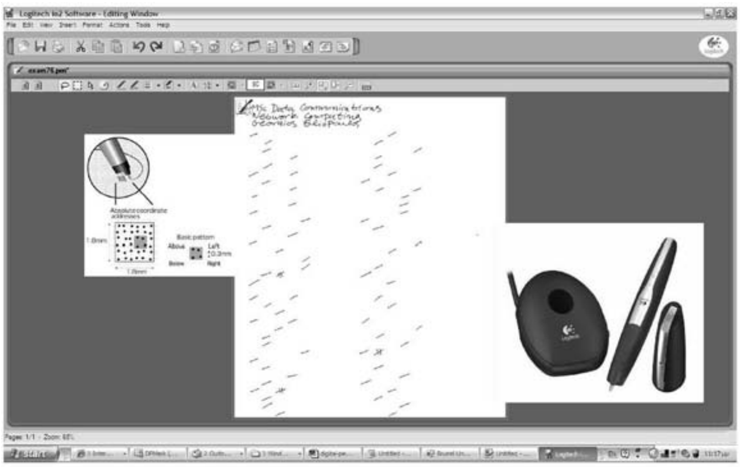
Abstract: Digital pens have been introduced over the last six years and have demonstrated that they can be used effectively for collecting, processing and storing data. These properties make them ideal for use in education, particularly in the marking procedure of multiple-choice questions (MCQ). In this report, we present a system that was designed to expedite the marking procedure of MCQ, for use at any educational level. The main element of the system is a digital pen, i.e. given to the students prior to the examination. On return of the pen, the system immediately recognises the students' answers and produces their results. In this specific research, four groups of students were studied and a variety of data were collected, concerning issues, such as accuracy, time gained by the use of the system and the impressions of the students. The pedagogic value of the use of the system is also presented.

Keywords: digital pen; marking; examinations; multiple-choice; expedite results; student satisfaction.

1. Introduction

Over the last six years digital pens have been introduced to the general public as devices that would be useful to the average person, offering a wide range of capabilities, such as keeping notes and later converting them to e-format, automatically recognising handwritten text, creating drawings and instantly distributing them among a group, etc. (Mascord et al., 2008). Although the introduction of the web and e-mail might have contributed to the achievement of the 'paperless office', in reality paper use has increased by an average of 40% in large organisations over the last eight years (Robinson et al., 1997). People insist on using paper due to its physical properties (thin, light and flexible) as these make it easy to read, carry, fold and write upon. Also many people do not wish to get distracted by the technology used in e-storage media such as PDAs, smartphones, personal computers (PCs), etc. (Bridgeman, 1992). The concept of the digital pen serves to bridge the gap between digital storage with all its advantages (physical space, indexing, retrieval and capacity) and physical paper, i.e. a more natural medium for written communication. Digital pens, much like ordinary pens, are light and portable, requiring very little power to operate while having a variety of means for connection to a computer, both wired (by placing the pen in its cradle) and wireless (usually using Bluetooth wireless system). A digital pen consists of an ordinary pen with the addition of a miniature camera close to the pen's tip and an advanced image processing unit. To operate, digital pens are used in conjunction with specially designed paper that includes microscopic grey dots, almost invisible at a reading distance. These dots are arranged in a pseudo-random pattern with different spaces between them. Each time the pen's tip scans over these dots, their pattern is read by the pen's camera and immediately recognised by the CPU of the pen. The change in the pattern indicates the specific route that the pen took as the image processor calculates the coordinates corresponding to each position in the entire pattern area. The series of points captured at 100 frames per sec is consequently converted to a line drawing, i.e. stored in the pen's memory for later retrieval (Figure 1).

Figure 1: Sample of an exam answer sheet downloaded and viewed on Logitech software



Note: The pen file (middle) will then be passed to the marking application.

The final step includes the pen communicating with a PC (either wired or wirelessly) and dumping the recorded information in the form of an '.xml' document containing both the attributes of the document (size, etc.) and the set of every pen stroke on the paper. This can then be manipulated using conventional means, stored, e-mailed, distributed, etc. The overall procedure is very accurate as the pen, and especially the dot arrangement in the paper, is specifically designed not to miss or misinterpret any information. A portion of 1.8 × 1.8 mm uniquely identifies the absolute pen position on a virtual canvas which covers an area exceeding 4.6 × 106 km2 (or 73 × 1032 A4-size pages), corresponding to half the Earth's surface area (Robinson et al., 1997). The above properties make digital pens ideal for use in various everyday life applications, such as in hospitals to fill out forms, in industry to distribute e-notes and maintain a paperless office, for elderly people who lack the ability to operate a computer effectively and, most importantly, in surveys. In the latter, the conventional method requires a person to make a series of 'ticks' on a form and manual feeding the results into a PC for statistical analysis. Automated scanners are available in the market and are designed to perform a similar task using a combination of electrical and mechanical means. However, as it will be demonstrated later, digital pens have considerable advantages over those scanners. Today digital pens seem to be under-utilised in environments like schools and universities where the demand for such applications is clear: from primary schools to the first university years, multiple-choice questionnaires have been successfully used for decades (Furukawa et al., 2004; Jassar, 2004). Current bibliography shows that the experts' opinion on the validity and effectiveness of multiple-choice questions (MCQ) seems to be divided with specific advantages and disadvantages being often mentioned (Kehoe, 1995). Some of the advantages include:

- high reliability and decreased chance of error while taking and grading the test
- less time spend by the student for answering the questions of each section, therefore, more sections can be included within a single examination thus testing a wider portion of the examinable material
- lowering the chance of subjective interpretation of the questions by the students
- producing less formatting mistakes as they only allow for one specific answer.

On the other hand, a number of disadvantages are also mentioned (Schouller, 1998):

- MCQs allow for pure guesswork while students are still able to get a substantial grade (statistically speaking, choosing random answers on an A/B/C/D MCQ will result in 25% correct answers).
- MCQs cannot account for students having partial knowledge of the question.
- Essays seem to be more suitable to test higher order skills (Hart, 2008).
- There can be uncertainly if the exam is not set correctly and more than one answer seems to be on the correct side.

Nevertheless, the value of multiple-choice assessments has become increasingly popular as class sizes have grown, especially, in higher education (Presser, 2004). Combined with the practical imperative to reduce the amount of time spent in assessing and providing feedback on work, the pedagogic arguments for their use have become clearer in two ways. Firstly, the analysis of learning outcomes, while, generally, taken to emphasise the importance of higher-level activities, such as evaluation and synthesis as the ultimate goals of higher education, also recognises the importance of earlier stages, such as the knowledge of facts. The systematic development of learning steps and assessment schemes that incorporate all levels of learning outcomes is a design feature informed by pedagogic principles relating to the setting of explicit subgoals and the motivational effects of success. Secondly, the early criticism of multiple-choice assessments – that they could only test factual knowledge – is no longer sustainable (Schouller, 1998; Schouller and Prosser, 1994). Work by Bush (1999) and Breland and Gaynor (1979) demonstrates that it is possible to devise such questions which require higher levels of thinking over and above recall or (worse) recognition. At the same time, it is possible to avoid obstacles posed by language unfamiliarity and disabilities. Results of students tested in both forms; direct (essay writing) and indirect (multiple-choice questionnaires) were found to be comparable (Bridgeman, 1992).

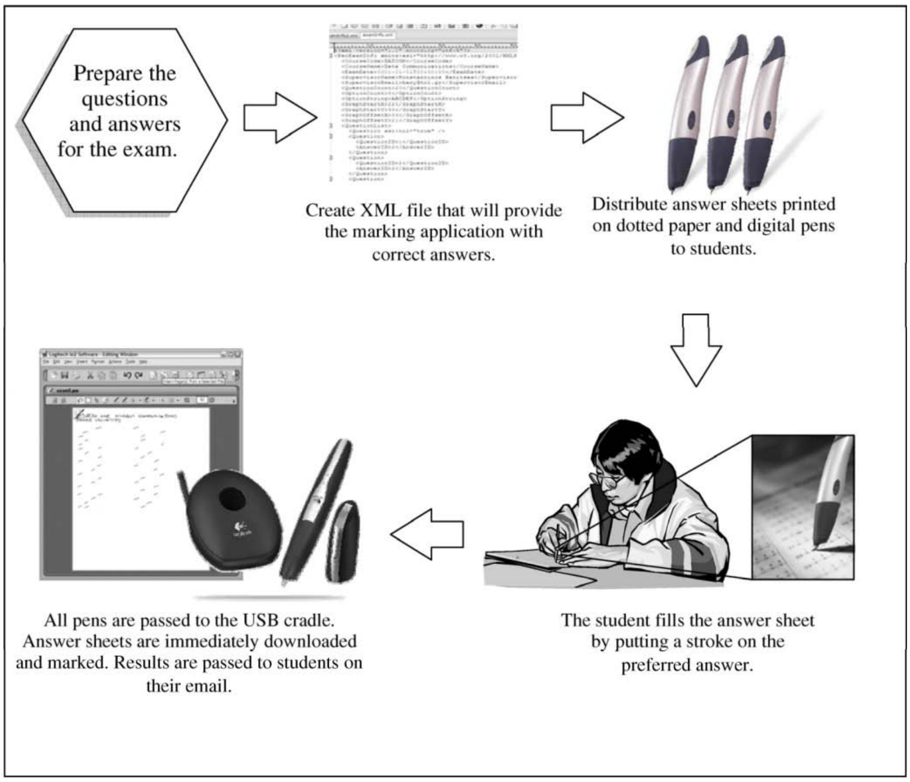
Furthermore, an essential aspect of ensuring that assessment supports learning (as well as providing the means to judge it) is the provision of effective feedback. A major characteristic of effective feedback is that it is timely; given soon after the assessment has been completed. Any system of assessment that speeds up marking with a concomitant reduction in delay in providing feedback to students is to be welcomed for sound pedagogic reasons (NSS, 2008).

Obviously, not all disciplines benefit the same by the use of such techniques. MCQs were found to be more effective in areas that demand a more surface rather than a deeper strategy. In a sample of 248 university students, e.g., Smith and Miller (2005) found that psychology students were more prone to use essay methods of describing a concept rather than MCQs. Alternatively, business students in the same sample preferred MCQs as they seem to relate their thoughts easier. In general, Roediger and Marsh (2005) list the possible positive and negative consequences of a MCQ test and argue that disciplines such as engineering, medicine, business and certain language courses can more effectively use MCQ to their advantage rather than disciplines such as psychology, human studies and education.

By taking into account the above-mentioned technology, combined with the use of multiple-choice assessments in educational programmes, one can foresee an obvious application for using digital pens to assist with the marking of MCQ at any educational level. Ideally, the system would be highly accurate, should produce results instantly with the possibility of forwarding those results to the student, and would save time for both the marker and the student and finally, be easy to use.

The prototype system used in this study consists of a Logitech 'Io2' digital pen that connects to a PC through a cradle. The PC could accommodate any number of digital pens and through specially designed software reads, records and forwards the results when the pens are returned to the cradle (Figure 2). The required specification for the PC is quite minimal. It is envisaged that an examination room could easily be equipped with a sufficient number of digital pens in order to conduct a multiple-choice examination, independent of any other technology (multiple PCs, scanners, internet, etc.).

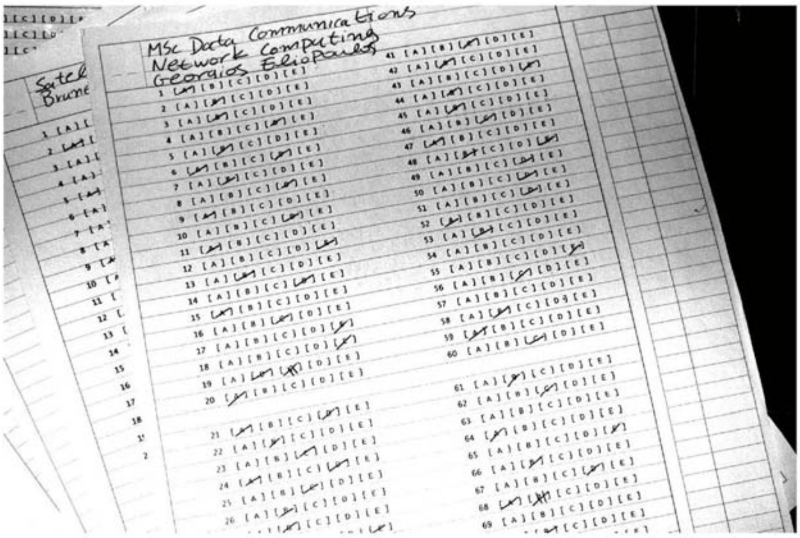
Figure 2: The digital pen marking process in five steps



2. Methodology

For this study, we used a Logitech Io2 that connects to a PC through a Universal Serial Bus (USB) port. We designed an examination paper having 80 questions which fitted onto a single sheet of paper. The paper was specifically designed for this Logitech pen and was readily obtainable from various sources. On this paper, a set of 80 answer areas (each having five square boxes representing answers A, B, C, D or E) were printed using either a jet or a laser printer (Figure 3). This method has previously been tested for any number of answers per page and also for more than one page per exam (Bastéa-Forte et al., 2007; Kolberg and Magill, 2006). In this study, the test addressed first year university students, although the same principles apply for any educational level (Duder and White, 1971). Students were given the MCQ on a normal piece of paper and were asked to indicate the correct answer on the specially designed answer sheet. No further instructions were given as to how they should proceed other than to indicate that if they checked on a wrong box, they should indicate their correct choice by checking another box and crossing out (put more strokes) in the wrong choice (Figure 3). This allowed the identifying software to decide the correct answer based on which one had fewer strokes inside the box.

Figure 3: Sample answer sheets used in the marking study



Note: They were designed using common word processing software and printed using special paper.

Four different groups of students (50 in total) took part in this study. At the end of the test, the students returned the pen to the operator who put it in the PC cradle; a procedure that took about the same time as students returning an examination script. Specially designed software based on Microsoft C# utilising the Logitech's Software Development Kit for the .NET framework was running on the host PC and automatically read the pen's data, recognised the responses, compared them with the correct answers and produced the results for each student. This file was then directly imported to an MS-Excel spreadsheet for further processing by the academic staff. Finally, feedback on the overall procedure was obtained from the students. The questions on which they were asked to comment included: ease of use, problems faced, overall feeling compared to conventional methods and any distraction felt as a result of the new technology.

3. Design elements

A communication protocol had to be defined for transferring the examination data (usually stored on paper or other software) to the marking application. Following the guidelines of modular programming, we initially introduced a blueprint of the information that needs to be transferred between the different modules and chose XML as a standard for object serialisation. Two basic objects ExamInfo (examination definition file) and ExamList (participants definition file) were identified as the source of data needed to initiate the marking application (Figure 4). ExamInfo can be considered as a digital version of the actual exam paper. It keeps the course code, name, examination date and the correct answers for each question. Alignment coordinates are also provided (GraphStartX, GraphStartY, GraphOffsetX, GraphOffsetY) and were used by the image analysing algorithm to match student answers to questions.

Figure 4: The ExamInfo XML serialised object


```xml
<?xml version="1.0" encoding="utf-8"?>
<PenExamInfo>
  <CourseCode>DATCOM</CourseCode>
  <CourseName>Data Communications</CourseName>
  <ExamDate>2008-01-14 15:00:00</ExamDate>
  <SupervisorName>Konstantinos Banitsas</SupervisorName>
  <SupervisorEmail>Konstantinos.Banitsas@Brunel.ac.uk</SupervisorEmail>
  <QuestionCount>20</QuestionCount>
  <OptionCount>5</OptionCount>
  <OptionString>ABCDEF</OptionString>
  <GraphStartX>22</GraphStartX>
  <GraphStartY>94</GraphStartY>
  <GraphOffsetX>34</GraphOffsetX>
  <GraphOffsetY>21</GraphOffsetY>
  <QuestionList>
    <Question>
      <QuestionID>1</QuestionID>
      <AnswerID>2</AnswerID>
    </Question>
  </QuestionList>
</PenExamInfo>
```


Figure 5: The ExamList XML serialised object


```xml
<ExamList>
<StudentList>
<Student>
<FullName>Georgios Eliopoulos</FullName>
<StudentNumber>0631232</StudentNumber>
<PenId>AR5-AAF-65P-MC</PenId>
</Student>
<Student>
<FullName>Chrysostomos Baxevanidis</FullName>
<StudentNumber>0631233</StudentNumber>
<PenId>AR5-AAF-65P-MD</PenId>
</Student>
</ExamList>
```


The second basic object ExamList provided a match of pen ids to each student using the pen (Figure 5). This file had to be generated just after handing all the pens to the students. The marking algorithm would then use this serialised list to match the downloaded stroke files to student numbers and produce the final scores.

To initiate the marking procedure there were two main sources of data: The ExamInfo and ExamList XML files that describe the questions and answers set along with the list of candidates and the downloaded pen files that were generated from the digital pens passing from the cradle. Then these have to be stored in appropriate objects that will provide interfaces to the application engine for data retrieval and parameter passing.

Following this two core functions were designed: the first is the one that will handle the exam data, read the correct answer to each question and retrieve the student names for every pen id number in the pen files. The second is the marking algorithm. Every pen file passed from the file system has to be analysed by this function stroke by stroke to produce the list of the candidate's answers and then compare those to the correct ones that are provided by the exam handler.

3.1. The marking algorithm

It was clear from the early steps of this project that coding the marking algorithm would be biggest engineering challenges. After considering the available options a decision was made to develop a new technique that would strongly be based on the Logitech Software Development Kit. An exam answer sheet was printed on Logitech special paper by simply enumerating the question from 1 to 80. On the right of each question number, we printed the available options separated by square brackets that would create some space necessary for the final recognition.

The final output follows:

```
 1 [A] [B] [C] [D] [E]    6  [A] [B] [C] [D] [E]
 2 [A] [B] [C] [D] [E]    7  [A] [B] [C] [D] [E]
 3 [A] [B] [C] [D] [E]    8  [A] [B] [C] [D] [E]
 4 [A] [B] [C] [D] [E]    9  [A] [B] [C] [D] [E]
 5 [A] [B] [C] [D] [E]    10 [A] [B] [C] [D] [E]
```

In the final template, we chose to add a space character on the left and the right of each option letter to allow more space for the input. This was decided to improve on the fact that small spaces found in existing optical readers require concentration on the paper and increase the probability the input error if the student runs out of time and is rushing.

One more advantage of this format is that it allows layout changes or number of questions due to its simplicity. The final setup came with a fixed 80-questions form printed in two columns on the same page but provision is made to accommodate different number of questions (this will be explained further in the following chapters). The next step was to gather input data from the pen file and overlay those onto a virtual grid, i.e. constructed by covering each option with a virtual rectangle. The marking process has assigned each possible answer to a rectangle area by reading the calibrated four coordinates from the ExamInfo XML file. An example of this follows:

```
<GraphStartX>22</GraphStartX>

<GraphStartY>94</GraphStartY>

<GraphOffsetX>34</GraphOffsetX>

<GraphOffsetY>22</GraphOffsetY>
```

The above code describes a grid with the upper left corner on coordinates (x = 22, y = 94) and each rectangle being 34 units wide by 22 units high. The algorithm also knows the number of options (5) and the number of questions (80) so the entire grid is constructed in an object structure where all grid cells are stored as rectangle object.

The marking algorithm will now parse through the stroke list on the pen file object (Figure 6). For each stroke, a path is constructed and the exact rectangle area surrounding this path is calculated. Then the intersection of this area and the stroke area is calculated to produce a hit or miss reply. C# provides a generic function for intersecting rectangles that accepts two input object and returns the area of the intersection (or zero if they do not intersect).

Each time a result is found, the algorithm will populate a summary list (Figure 7).
The fields displayed on the first line from left to right are:

Stroke57 = The identification string of the examined pen stroke.

10/1/2008 8:01:24 = Timestamp of the pen stroke.

Q1 = The question number on which the stroke was marked.

E = The option letter selected by the candidate.

25 = The number of point on the stroke path.

288,5557 = The area result of the intersection function

Figure 6: The virtual grid in the marking algorithm

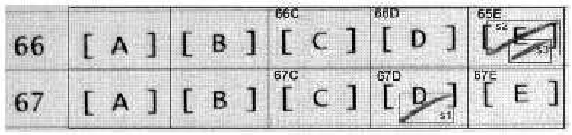
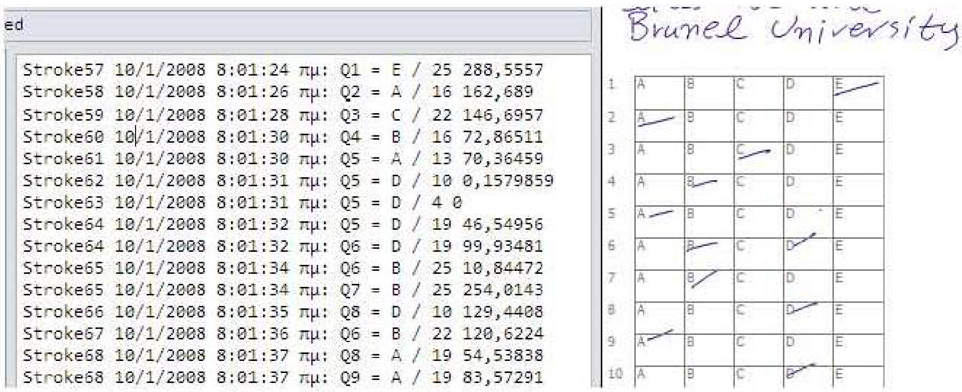


Figure 7: Elements of the marking algorithm interface (see online version for colours)



The key values in identifying accidental pen strokes are the number of point per stroke and the area size of the recognised stroke. In Figure 7/Question 5 on the right, we demonstrate a case when there is more than one stroke identified inside the answer boxes area The algorithm can correctly identify answer A as the entered value by comparing the area and number of point between candidate strokes (candidate A = 13 points/70 units area size against candidates in box D).

4. Results

The results were analysed in terms of accuracy, time, security and user feedback.
Accuracy and time taken were compared with manual marking of the examination.

4.1. Accuracy

The overall accuracy of the system was 100%. Over the sample of 50 answer sheets that were handed in by students, there were no errors and no discrepancy between a student's answers and the computer's recognition of the answer. The examination answer sheets were also marked by conventional means so that the digital pen accuracy could be compared with manual accuracy. In this case, on 50 scripts with 80 questions each, there were a total of 4 initially misread answers by the person marking the scripts manually.

The situation can be improved by printing the model answers on a transparency and placing them on top of the exam to make right/wrong identification easier and faster. In general, the manual marking of MCQ results in an accuracy of about 98–100% (Presser, 2004).

4.2. Required time

The time needed from the moment the student placed the digital pen into a cradle until the moment the results were produced was about 1–2 sec using an average PC. This time would be considerably longer using manual means. Although no strict record of this was taken, it was estimated that an 80-question examination would take a minimum of 4 min to mark and would require the full attention of the marker.

4.3. Security

In theory, the security of the system should not be an issue as, in our study, we used wired rather than wireless technology. This means that the pen uniquely identifies itself and cannot be duplicated or tampered with. The possibility of students swapping digital pens between them is dealt with using various solutions described below.

4.4. User feedback

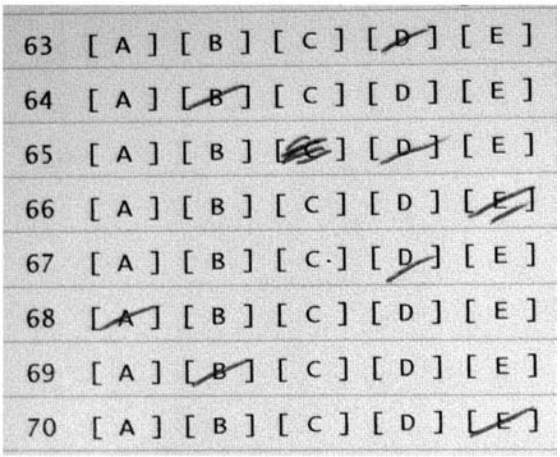
All students participating in this test agreed that it was an interesting, original and helpful application. Most of these students had an engineering background. They were mainly interested in having their results returned as soon as possible and thus avoiding long waiting periods of uncertainty. As indicated above, delay in the provision of feedback is a common and ongoing complaint of higher education students (Hogan, 2007; NSS, 2008, q.7). A large majority of students reported that the system was easy to use (94%). When asked for explanation of their negative response, the remaining 6% gave a strong indication that they felt somewhat alienated towards this kind of technology and compared it with other hi-tech modern gadgets (mp3s, smartphones, etc). Students also reported that they had no major issues with the technology behind it or with its use (96%) and that they had a positive feeling towards this technology as it would expedite the production of their grades (100%). Upon the matter of how much they trusted the fairness and accuracy of the automated marking, 99% answered that they would not consider appealing with regard to such a procedure. Finally, they reported that, although at the beginning they felt a bit distracted by the technology, they soon regained their focus and were able to concentrate on the exam questions (100%).

Three problems had emerged during a pilot procedure and these had been addressed as discussed below.

4.5. Change of mind

Sometimes students change their minds about what the right answer is and tick a second answer, smudging their original choice. This creates a problem even when using conventional marking and it certainly created a problem during the pilot stage for this study. It was found that it would be easier to instruct the students to mark any wrong choice with *more* strokes (cross it out) than their final answer. The software was then developed to compare each possible answer box and disregard the answer(s) that were bolder (Figure 8, line 65).

Figure 8: For question 65 ‘D’ is the correct answer

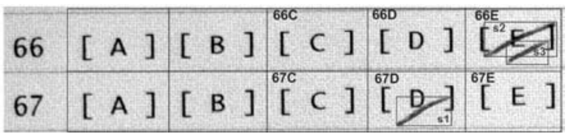


Note: ‘C’ is considered as deleted after four stokes were detected. The dot on 67-C is dismissed as the rectangle area around the stroke is below threshold value.

4.6. Accidental writing

*S*ometimes students playing with the pen left a tiny mark, a dot, on one of the boxes without realising it. To deal with these potential problems, the software created a threshold level below which no answer is taken into account (a single dot would be dismissed). (Figure 8, line 67C and Figure 9, line 67).

Figure 9: The intersection of rectangles: the marking process has assigned each possible answer to a rectangle area



Note: The intersection of this area to the stroke area is calculated to produce a hit or miss response.

4.7. Handwritten text

Digital pens also have the ability to read and convert handwriting into text. This could have been, particularly, useful as the student names and numbers could have been recorded automatically without any human intervention. However, previous experience has shown that due to the wide variety of hand writing styles, the outputs would suffer from great inaccuracies. In our research, the option of recognising handwritten text was not used. Instead, the operator of the system would either manually log the results to a specific student or have each student assigned with a specific digital pen's code, before commencing the examination.

5. Discussion

This research examined whether using the proposed system would expedite the marking procedure, while maintaining (or even improving) the accuracy, compared with using conventional marking methods and whether combining this technology with the use of MCQs would have any pedagogic advantage.

Through the results described above, it was clear that the system would be beneficial as it reduces the processing time of an examination considerably while increasing the accuracy by avoiding common causes of human error such as fatigue induced by large marking loads. Even though the system performed accurately, it is recommended that, at least in the first stages of its potential rollout, there should be some random sampling of exam scripts accompanied by manual marking to ensure reliability. Furthermore, in order to increase security, the student should be required to write his/her name and student number at the top of the page (Figure 3). This would ensure that in the worst-case scenario, the exam can be marked manually while having all the necessary information on the same page. It will also act as a fail-safe mechanism in the case that students might switch digital pens between them: when the pen is returned to the cradle, the handwritten text (student's name and number) will be visible on the operator's screen for verification purposes.

As mentioned earlier, the developed software also contains a web-based application that would allow student to directly login and get their results (either as a total score or by having each of their answers indicated as 'correct' or 'incorrect'). The latter option provides some level of feedback, which is much appreciated by students.

During the latest stage of the system development, two more modes of feedback were developed: the first takes into account the student's answer and in case it is wrong, provides the student with the correct answer and the relevant section of the book/teaching notes that demonstrate the correct answer (e.g. a paragraph of the textbook that directly relates to the question set in the exam sheet). This, however, assumes that the teacher/lecturer has prepared the appropriate sample answers for each question. The second is, somewhat, more complicated and identifies areas of weakness to the student's answers. The software looks for incorrect answers to questions within the same learning area and provides a more conventional type of feedback. For example, in a mathematics exam the software will run a routine like:

IF q23 NOT EQUAL correct AND q32 NOT EQUAL correct AND q47...THEN DIPLAY comment

If this routine identifies that the wrong answers are mostly in the geometry section, it can automatically create a comment like 'There seems to be a weakness in the area of geometry. Please revise pages 45–52 from the textbook'.

Finally, these results can be produced in such a format that would allow direct input into the student record without any human intervention throughout the whole procedure (Figure 11). However, there might be administration conflicts with this option as, usually, in higher educational levels, exam results have to be validated by an Examination Board before being formally recorded.

Figure 10: The user interface of the marking software: by clicking on the student list on the left-hand side the application retrieves the relevant pen file, loads it on the right-hand side frame and executes the marking algorithm

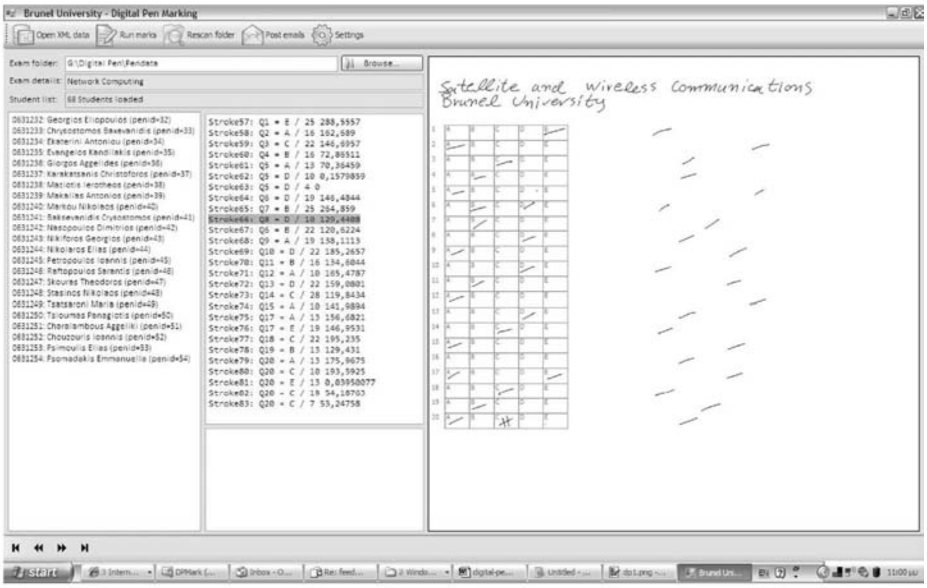
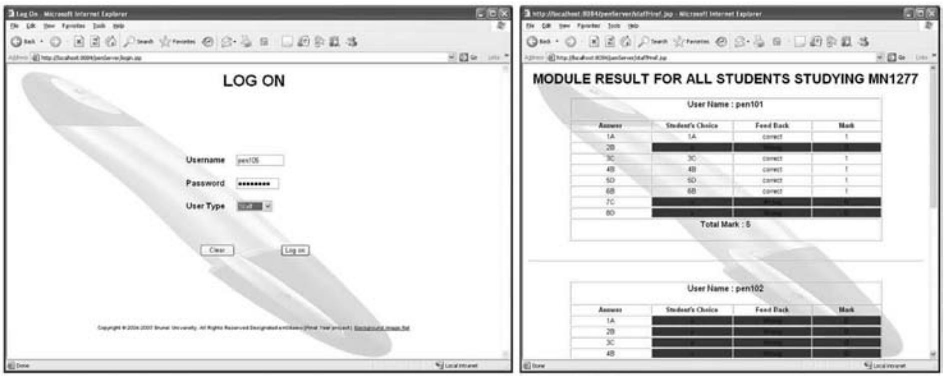


Figure 11: Exam results can be directly published on existing e-learning applications



The system can be deployed in a large-scale classroom (irrespective to the number of students) needing one invigilator and a cradle connected to a single PC. An all-wireless system based on Bluetooth digital pens can also be developed by slightly modifying the software. Students can wirelessly transmit their result while still sitting at their desks during the exam. This solution, however, might lead to networking problems (there would be some level of interference between the digital pens that would limit their nominal range of about 10 m radius) or even loss of property; a problem not faced in the wired method as the student has to return the pen in order to log his/her results.

For the wired solution, and as the number of students increases to more than 15–20, more than one computer should be available to avoid long queues after the exam. All computers will be feeding the data to the same database so there is no limitation as to the number of wired cradles/computers to be used. The scenario differs for the wireless (Bluetooth) solution as the many-to-one network established (many digital pens to one receiving station) is self-configurable: Bluetooth networks only allow one transmission at a time, putting the rest of the stations in a queue. As the digital pen transmits data of an average exam sheet in a few milliseconds, it only takes a few seconds to complete the transmission for a class of 30 students (Figure 10).

Along with the technical advantages of the system, there are considerable pedagogic advantages:

- Students using the proposed system would receive an almost immediate feedback allowing them to restrategise their studying and fill in their educational gaps quicker. Moreover, if the set up of the exam allows for specific feedback to be released to the student in the form of 'there seems to be a weakness in the area of...', the student can quickly and timely concentrate on that area.
- Timely feedback also leads to increased student satisfaction as the student feels as being 'looked after' by the educational organisation. This increases the chances of the student reciprocating the gesture thus behaving more professionally towards deadlines. The objective is for the students to feel that if the educational organisation went to that length to provide them with quick and accurate feedback; it would only be fair to be prompt and punctual themselves.
- Furthermore, since the system is automatic and fully computerised, the possibility of error is very close to zero. This will lead to students feeling less anxious and doubtful towards the results and will eliminate any suspicion that an examination marker might have been biased towards them.
- Such a system, if properly designed, will decrease the chance of a student cheating during an exam by introducing objects like already filled-out pieces of paper, pens/pencils/rulers that might include forbidden material. The system only allows for specific papers and pens to be used without the need for any other objects to be present in the examination room. Students, therefore, are more likely to concentrate on the exam itself using the best of their abilities while being less tempted to waste time experimenting with ways to cheat.
- Finally, valuable and timely feedback can also be given to the teacher/lecturer, instantly identifying areas that most students did not do well. A post-processing of all the results yields for a three-dimensional table to be created (question number x choice x student number). Mistakes appearing with increased frequency yield for a topic that was probably not sufficiently understood by the class.

A final thing to mention is that automatic optical readers are not something new, as they have been used for decades. However, when compared with existing technologies, digital pens outperform both the use of optical readers and online multiple-choice tests. Optical mark readers have, traditionally, been used to automate the marking procedure of multiple-choice questionnaires; in particular, this has been the case in the public sector including the army and large commercial organisations, such as banks (Haag et al., 2006). The system investigated here ranks at a similar level of ease and accuracy, compared to the optical reader. Both have an initial cost of about £2,000 and produce results with similar accuracy and speed. Digital pens, however, are considered more flexible as they support open-source software that allows for a faster reconfiguration of an examination format (number of questions, exam layout, inclusion of figures and schematics to be completed by the student). They are also more secure (pens are student-aware since, during the exam, every digital pen's unique ID corresponds to a student's ID) and allow for a change of mind of the student. Optical readers require the student to use an eraser which has a high potential for creating false-positives.

Online tests, on the other hand, require a class with a number of PCs matching the number of students. It is clear that the cost of each PC is considerably higher than the cost of a digital pen. The availability of large examination rooms equipped with PCs for each candidate is not widespread due to the high cost of provision. Furthermore, there are security considerations as there could be information stored or available for access within this PC to which students should not have access during the examination. Finally, using a traditional method of pen and paper tends to distract the student less than following the instructions of custom designed software on a PC screen.

The cost of the digital pen system is quite modest. Each digital pen, along with the cradle costs about £60–70. Any low-end networked PC having a USB port can be used as a host. Thus the overall cost of the system deployed on a 30 student class would be around £2,000–2,500. One lecturer would take an average of 2–3 hr to mark an 80-question multiple-choice-based exam for a class of 30 or more students. It is, therefore, anticipated that the system will cover its expense after about 25 cycles of exam marking.

6. Conclusion

This study has demonstrated the use of a technologically enhanced system for automatically marking multiple-choice-based exams. Investigation of the system in practice showed that it reduced the time needed to mark examination answers to almost zero, while having almost 100% accuracy (higher than expected from a person performing a similar task). The system also demonstrates some advantages in pedagogic and didactic aspects, desirable both to the students and the teachers. Feedback from the students also showed that the system had attractive properties (friendly, easy to use) and its technology did not distract the user. Its cost is relatively low and it can compensate for its expense in a short period of time.

Overall, the system's properties make it a desirable tool when used where multiple-choice examination can provide a pedagogically valid method of assessment.

References

Bastéa-Forte, M., Brandt, J. and Boli, J. (2007) *Iterative Design of a Paper + Digital Toolkit: Supporting Designing, Developing, and Debugging*. Stanford: Stanford University HCI Group.

Breland, H. and Gaynor, J. (1979) 'A comparison of direct and indirect assessment of writing skill', *Journal of Education Measurement*, Vol. 16, No. 2, pp.119–128.

Bridgeman, B. (1992) 'A comparison of quantitative questions in open-ended and multiple choice formats', *Journal of Educational Measurement*, Vol. 29, No. 3, pp.252–271.

Bush, M. (1999) 'Alternative marking schemes for on-line multiple choice tests', *7th Annual Conference on the Teaching of Computing*, Belfast.

Duder, J.C. and White, G.R. (1971) 'A reliable multiple choice test marking system', *Physics Education*, Vol. 6, No. 3, pp.165–166.

Furukawa, N., Ikeda, H., Kato, Y. and Sako, H. (2004) 'D-Pen: 'a digital pen system for public and business enterprises', *Ninth International Workshop on Frontiers in Handwriting Recognition*, 26–29 October, Tokyo, Japan, pp.269–274.

Haag, S., Cummings, M., McCubbrey, D., Pinsonnault, A. and Donovan, R. (2006) *Management Information Systems for the Information Age* (3rd ed.). Canada: McGraw-Hill Ryerson.

Hart, P. (2008) 'How should colleges assess and improve student learning? Employers' views on the accountability challenge', *Association of American Colleges and Universities Report*.

Hogan, T. (2007) *Educational Assessment*. NY: Jossey-Bass, pp.152–174.

Jassar, S. (2004) 'The digital pen as a human computer interface', *IEEE International Symposium on Consumer Electronics*, Reading, UK, pp.629–634.

Kehoe, J. (1995) 'Writing multiple-choice test items', *Practical Assessment, Research and Evaluation*, Vol. 4, No. 9.

Kolberg, M. and Magill, E. (2006) 'Using pen and paper to control networked appliances', *Communications Magazine, IEEE*, Vol. 44, No. 11, pp.148–154.

Mascord, M., Jirotka, M. and Carusi, A. (2008) *Putting Pen to Paper: Capturing the Integrative Biologist's Paper Trail*. Oxford University Computing Laboratory, Oxford e-Research Centre.

National Student Survey (2008) *NSS Questionnaire*. HEFCE. Available at: http://www. thestudentsurvey.com/archive/nss-questionnaire-english.pdf. Accessed on October 2008.

Presser, S. (2004) *Methods for Testing and Evaluating Survey Questionnaires*. NY: John Wiley & Sons Inc, pp.50–65.

Robinson, P., Sheppard, D., Watts, R., Harding, R. and Lay, S. (1997) *A Framework for Interacting with Paper*. University of Cambridge, Computer Laboratory.

Roediger, H. and Marsh, E. (2005) 'The positive and negative consequences of multiple-choice testing', *Journal of Experimental Psychology*, Vol. 31, No. 5, pp.1155–1159.

Schouller, K. (1998) 'The influence of assessment method on students' learning approaches: multiple choice question examination versus assignment essay', *Higher Education*, Vol. 35, No. 4, pp.453–472.

Schouller, K. and Prosser, M. (1994) 'Students' experiences in studying for multiple choice question examinations', *Studies in Higher Education*, Vol. 19, No. 3, pp.267–279.

Smith, S.N. and Miller, R. (2005) 'Learning approaches: examination type, discipline of study, and gender', *Educational Psychology*, Vol. 25, No. 1, pp.43–53.