Remote electronic examinations: student experiences

Pete Thomas, Blaine Price, Carina Paine and Mike Richards

All of the authors are members of the Centre for Informatics Education Research (CIER), in the Department of Computing, Open University, Milton Keynes, UK. Email: p.g.thomas@open.ac.uk

Abstract
This paper presents the findings from a small-scale experiment investigating the presentation of a synchronous remote electronic examination. It discusses the students' experiences of taking such an examination. The study confirms that the majority of participants found the experience at least as good as a conventional written examination. In addition, typing answers does not prevent students from producing answers in the time available. However, the pressure of time continues to be a major cause of anxiety for students. The paper discusses technical issues, particularly those related to the loss of communications during the 3-hour duration of the exam. Although software processes were available to save and restore students' answers throughout the examination, problems still occurred and more robust software is required.

Introduction
There is an increasing interest in the use of computer-assisted assessment (CAA) in higher education that is primarily focused on objective testing in assignments (Bull and McKenna, 2000). Much less has been reported about techniques for invigilated examinations despite the extensive reliance on them for summative assessment (O'Hare, 2001), although moves are being made to use software in public examinations (Long, 2000). However, there are many issues that need investigation including student anxiety (Carswell, Thomas et al., 2000; McKenna, 2001), automatic grading of free form answers (Burstein, Leacock et al., 2001; Shermis, Mzumara et al., 2001) and the reuse of questions (Bull and McKenna, 2000).

A remote electronic examination is presented to candidates at a location remote from the examining institution using the Internet for communication. Candidates respond to questions by typing their answers into text boxes for uploading to the institution. In an asynchronous examination candidates download the exam paper from a web site, prepare their answers off-line, and reconnect to the examination web site at the end of the time period. In a synchronous examination candidates remain connected to a server for the duration of the examination period. Our experiences with asynchronous
examinations are reported in (Thomas, Price et al., 2001) and in this paper we describe the students’ experiences of a synchronous examination.

With the growth in numbers of candidates to be examined, the prospect of grading examinations automatically promises faster, cheaper and more consistent grading (Shermis, Mzumara et al., 2001). Even if automatic marking is not used, capturing candidates’ answers electronically has potential benefits in legibility and comprehension for graders. There are advantages, too, in security, with papers being held electronically and only being released to candidates shortly before the designated start time of the examination. In a distributed system, as commonly found in distance education, electronic examinations have the potential for speeding up the whole examination process from the transfer of student answers to markers, performing standardisation and agreeing the final grades. One approach, adopted by Cambridge Examination Board in relation to their activities in Singapore, scans students’ handwritten answers and transmits them as bit maps to markers who use a software marking tool to grade and record marks (Harding, 2002). In our environment we want to see the extent to which the marking time can be reduced by the adoption of automatic marking. This significantly changes the student experience of examinations from a written to a typed exam and it is important to understand the impact of such a change. We can assess the impact in two ways: by investigating students’ experiences of the new medium and by comparing performance in the two environments. In this paper we concentrate on student reactions to working in the new medium.

A common criticism of remote examinations used for summative purposes is the difficulty of ensuring that cheating is minimised (Whittington, 1999; Frohlich, 2000). Electronic examinations taken under supervised conditions have been implemented (O’Hare, 2001), but we are interested in pursuing the use of examinations in less formal settings, particularly in the home. Such environments are similar to those in which distance education students normally study and avoid the need to attend unfamiliar locations that only heightens student anxiety. There are circumstances in which unsupervised examinations are quite acceptable, for example, giving students experience of the examination process through mock (pretend) examinations.

A synchronous examination
In the spring of 2001, students on a part-time, post-graduate, distance education course in computing were invited to volunteer to take a synchronous mock examination. Sitting the exam would provide experience of a formal, timed examination with questions similar to those on a real exam. The volunteers registered their interest by providing a password that would eventually give them access to the mock exam. Students were told to contact our server up to half an hour prior to the published start time of the examination in order to ensure that they could overcome any connection difficulties. We employed Microsoft NetMeeting for student-to-invigilator communication so that only one phone connection would be needed for both examination and problem resolution purposes. If problems were to arise during the timed period, students were told to contact an invigilator using either typed chat or audio. All
participants took the exam at home using a modem to connect to their own ISP. (The students were familiar with using computer communications for submitting coursework assignments (Thomas, Carwell et al. 1998).)

At the appointed start time the examination paper was made available for download (a message was sent to all participants informing them that the paper could be accessed). Initial access to the paper was restricted to the first half-hour of the examination period to simulate what might happen in a conventional written examination. Students could submit their final answers thereafter. In the event, all students started at about the official start time and submitted their answers towards the end of the three-hour period.

The exam paper was displayed in a web browser as illustrated by the two screen-shots in Figure 1. The design of the paper was rudimentary with an initial paragraph setting out the rules (rubric) for completing the exam, followed by a sequence of exam questions and text boxes for answers. There was an index to the start of each question at the head of the exam paper and a button alongside each answer for navigation purposes. The whole paper was scrollable.

Having submitted their answers, students received a thank-you receipt and an invitation to fill in a short questionnaire about their experiences. Ten students completed the questionnaire and their responses are discussed in the next section. The students’ answers were marked automatically and the results returned the next day. The results of this part of the experiment are discussed in (Thomas, Price et al. 2001; Thomas and Paine, 2002). The answers were also marked by two human graders. An average grader mark together with feedback on performance was sent via email within a week of the examination and students were surveyed once again about their experiences after reflection on the process.

Feedback
Immediate feedback
Figure 2 shows the responses to the question, “How do you rate this experience compared to a normal examination?”

To see what aspects of the electronic examination led to an inferior experience for students we first looked at the students’ immediate reactions, grouped by experience category and listed in Table 1.

If we divide the students into two groups, (1) those that found the experience about the same as or slightly better than a conventional examination, and (2) those that found the experience slightly or much worse, we find that there are issues common to both groups (but not necessarily reported by all students):

• a difficulty with periodically saving answers (particularly towards the end of the exam period);
• the lack of a facility to enter diagrams in answers;
• the mock exam was taken before they had completed their revision;
• a preference for writing over typing.
However, there are three issues mentioned by those who found the experience as slightly or much worse but not mentioned by the others and which may account, at least in part, for their poor reaction to the experience:

- it took a significant amount of time to become familiar (oriented) with the new situation (hence time was lost in answering the questions);

Figure 1: Screen shots from the synchronous examination paper

© British Educational Communications and Technology Agency, 2002.
a lack of confidence in the use of the electronic medium (mentioned by a self-styled incompetent user);

a loss of time at the start of the exam due to the system not accepting the password.

It seems clear that anything that might result in a loss of time (slow typing, orientation, gaining access to the exam, periodically saving answers) is a major issue for all students.

Figure 3 shows the responses to a question about the length of time to type in answers to questions, with Figure 4 showing the students’ own rating of their typing speed.

The students who reported having “not enough time” or “just enough time” to type their answers rated themselves as “hunt and peck” typists or “fluent 2–4 fingers”. The students

Table 1: Immediate student reactions

<table>
<thead>
<tr>
<th>Overall experience</th>
<th>No of students</th>
<th>Issues raised</th>
<th>No of times issue reported</th>
</tr>
</thead>
<tbody>
<tr>
<td>Slightly better</td>
<td>6</td>
<td>Difficulty with saving answers</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Lack of facility to enter diagrams</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Lack of revision</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Prefer typing to writing</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Prefer writing to typing</td>
<td>1</td>
</tr>
<tr>
<td>About the same</td>
<td>1</td>
<td>Difficulty with saving answers</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Lack of revision</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Artificial situation</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Prefer writing to typing</td>
<td>1</td>
</tr>
<tr>
<td>Slightly worse</td>
<td>1</td>
<td>Lack of diagrams</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Orientation takes time</td>
<td>1</td>
</tr>
<tr>
<td>Much worse</td>
<td>2</td>
<td>Difficulty with saving answers</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Lack of revision</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Prefer writing to typing</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Lack of confidence in sitting an exam</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Password problem at start</td>
<td>1</td>
</tr>
</tbody>
</table>
who reported having “plenty of time” to type their answers rated themselves as “fluent 2–4 fingers” or “fast 5–8 fingers”. Hence, students who are not proficient typists would seem to be at a disadvantage. However, of the three students who stated that they did not have enough time to complete the exam, one managed to attempt 100% of the required number of questions, and a second completed 95%. The third student was the weakest student who only managed to complete 73% of the required questions. However, another student who attempted only 80% of the required questions stated that she had “plenty of time” because she had completed as many questions as she was able as she had not completed her revision.

The lengths of answers from all students were quite short—certainly shorter than would normally be expected in a conventional written exam. Even so, answers were sufficiently comprehensive for marking purposes.

Table 2 shows the amount of typing performed by each student. The first row contains the total number of characters typed, the second row gives the number of characters per minute of the exam, and the third row shows the number of words typed per minute of the exam. The fourth row (entitled “Self-speed”) contains the students’ own estimates of their typing proficiency on the scale: 1 = “hunt and peck” to 4 = “touch
Typist”. If one ignores variations in the amount of time taken in reading and thinking, the students’ actual speeds closely match their own estimations.

The rows entitled “Enough time” and “Completed %” compare the students’ statement about whether or not they had enough time to complete the exam (on the scale: 1 = “not enough time”, 2 = “just enough time”, 3 = “plenty of time”) with the actual completion rate. The “Completed %” row gives the maximum number of marks (out of 100) that could have been obtained if the student had correctly answered those (parts of) questions they attempted.

Generally students were able to submit answers to the required number of questions in the time available, and where students felt that more time would have been useful, they had been able to complete a substantial proportion of the exam. This is in line with experience of conventional written examinations. We conclude, therefore, that typing is not necessarily a major impediment to sitting electronic examinations on a course where students regularly use computers for assignment work.

The “Experience” row is the students’ stated view of whether their experience of sitting the electronic examination was better or worse than their experience of sitting a conventional written examination under strictly invigilated conditions. The scale used was: 1 = “much worse”, 2 = “slightly worse”, 3 = “about the same”, 4 = “slightly better”, 5 = “much better”. The final two rows rank the group of students according to their performance on the mock exam (electronic) and real exam (written). Student 11 was the weakest student gaining a bare pass on both mock and real exams. All but one of the students performed better on the real exam than the mock.

Later feedback
Students were sent a second questionnaire along with the tutor marked copy of their examination script. All 11 students responded. Nine students said that they saw the mock exam as a revision opportunity and six said that they were interested in electronic assessment. The students believed that they were sitting an exam that fairly represented the course and did not find difficulty understanding the questions.
Generally, the instructions and screen design were acceptable (ensuring that text fits on
the screen is a well-known problem with current web browsers). The only problem with
logging-in was the use of a password. Two people had forgotten their password (we
implemented a “reminder” system that solved this problem); one person did not know
that a password was required, even though they had supplied one earlier; the system
refused to allow one person to log-in even though they were using the correct
password. Clearly, in any real system, passwords are going to be a problem for some
people and while reminder systems will help some, others will have difficulty.

Students were encouraged to save their work on a regular basis in case a problem with
their Internet connection arose. In the event, 8 of the 11 students had such a problem
and had to restore their answers. The majority of these students lost some work—
whatever they had entered between their last save and the time their Internet
connection failed. The one student who did not lose data had taken regular local copies
of her work. Whilst some problems with Internet connections is to be expected during
a continuous 3 hour period, we were surprised by the extent to which ISPs were the
root cause, placing limits on the time that their customers could stay continuously
on-line (although the situation has improved recently).

There has to be a system for recovery after a disaster especially given the frequency with
which Internet connections are lost. There was a bug in the save software that was fixed
on the day, but it was also clear that many of the candidates did not understand web
browser technology sufficiently for them to restore their work effectively. In most cases
students were able to restore their work after receiving help from an invigilator.
Considerable time can be lost if a recovery cannot be effected easily or saves have not
been done regularly. (The invigilators kept watch over saves, and reminded a candidate
to use the facility; although this is hardly a scalable activity. We envisage the use of
automated robots to solve this problem.)

Four students stated that they had problems navigating through the paper, particularly
having to scroll when they wanted to move randomly between questions despite the
existence of an index. Scrolling was seen to be time consuming. Four students printed
the exam paper in order to help them when making their selection of questions. These
points resulted in the redesign of the paper as shown in Figure 5. Five students reported
doing some rough work on paper during the examination. In most conventional
written examinations such work would have to be submitted along with their answers,
which is clearly not possible in an electronic setting.

Students were asked to list what they thought were the advantages and disadvantages
of an electronic examination and the results are shown in Table 3.

It was to be expected that, in a distance education environment, avoiding having to
tavel to an examination centre would be the one overriding advantage. The dis-
advantage most often reported was the lack of robustness in the technology (specifically
the loss of Internet connections and the problem in saving answers).
It is somewhat surprising that the issues of typing and lack of a diagramming facility were not reported more often as disadvantages given their prominence in the immediate feedback. The only factor that we can point to that might account for this outcome is that when the students completed the second questionnaire they had received a mark and feedback on their performance which may have allayed some of the fears previously expressed. None of the students had had any previous experience of an electronic examination.
All but one of the students felt that the exam was fair and that the questions were clear. The one student who thought that the questions were unfair (stated as “unexpected”) felt that the experience was slightly better than for a conventional exam!

The issues that students highlighted can be divided conveniently into four categories:

- **Technological**—related to the communications medium, access to paper, and printing;
- **Technical**—related to the format of the exam paper and the use of diagrams;
- **Environmental**—related to typing, time, artificiality, and orientation;
- **Personal**—related to revision, confidence in use of technology, and confidence in sitting exams.

Whilst the vagaries of the Internet and ISPs in particular are not within our control, we can certainly improve our software to make it more robust and hence improve the students’ confidence in the system as a whole. The format of the exam paper can be improved in the light of feedback and, given the extensive literature on pattern recognition for diagrams, we envisage being able to offer a diagramming facility for input.

Some improvements can be envisaged to some environmental factors. For example, orientation can be improved by providing a test paper that students can access prior to the mock exam.

However, the issues categorised as personal are more difficult to deal with. It is the nature of a mock exam that it will pre-date the final exam and will be taken before the majority of students have completed their revision.
Despite the volunteers on this trial being students of computing, some were lacking in confidence in the use of the technology as evidenced by the amount of help required to overcome difficulties with the technology.

The feeling that the experience was not like a real exam (“artificial”, “not a special occasion”, “no adrenaline rush”) reported by two students was not shared by the other students, seven of whom felt that they were sitting a real exam. Given that this was a mock exam with no summative component will have contributed to a feeling of artificiality. However, we hope to use the system to help students who have not taken formal examinations for some time to re-orient themselves and we are encouraged by the majority report that they did indeed feel as though they were sitting a real exam.

Despite the problems that arose, students volunteered the following comments that encourage us to continue with our investigations:

“Interesting experience!! Very well managed. I was impressed.”
“I enjoyed doing this.”
“Looking forward to see if this will be the exam of the future.”

Web-based examination papers
There are significant differences between the structure of an electronic exam paper and a paper-based exam paper. Clearly, there are issues related to fonts (choosing the appropriate type and size for the medium, for example), but there are other issues relating to usability and functionality to be considered. Figure 5 shows the structure of the latest version of our examination paper as it would appear in a candidate’s web browser (designed for synchronous use—an asynchronous paper would have a subset of these features).

The paper has a fixed border area giving access to commonly required functions, and an internal scrolling frame where questions and answer text areas appear.

The left-hand border contains a hyperlinked index giving single-click access to the start of each question. This index also displays indicators relating to the state of the answers; for example, whether or not each question has been attempted and whether the candidate has indicated their intention to return to the question at a later time. At the bottom of this border is a clock that indicates the time remaining.

The right-hand border has buttons for four tools:

Submit—used at the end of the exam to submit the answers to the server.
Mark unfinished—used as a reminder that the candidate aims to return to the current question at a later time.
Save—saves a copy of the current state of the answers on the server.
Retrieve—retrieves the state of the answers (at the point of the last Save) from the server.
The Save and Retrieve buttons are used to overcome the problems that can arise should the candidate’s Internet connection be lost. The current implementation of the paper requires the candidate to make regular saves, but we aim to include automatic saves in a future implementation.

Along the top of the border area are various buttons relating to downloading versions of the paper, including one for printing locally (required, for example, for candidates with special needs).

The internal scrolling area displays the questions and provides text areas in which to type answers. Currently, only simple textual input is allowed but we are working on prototypes for rich text and graphical input (Foxley, Higgins et al., 2001).

Future work
The feedback obtained from this small-scale experiment has enabled us to improve certain technical aspects of the process. Improving the robustness of the software and changing the design of the exam paper will remove some of the reported disadvantages of the process. This will enable us to concentrate on some deeper issues related to the student experience. Clearly, the results reported here cannot be said to be conclusive as they are based on a very small sample of students generally accustomed to using keyboards and sitting a mock rather than a real examination. We therefore are repeating the experiment with larger numbers of students to determine more precisely those features of a remote electronic exam that students find preferable and those that they dislike and the reasons for these views.

The issue of typing during a time-limited examination is a crucial factor in determining whether or not to adopt electronic examinations for summative use. The present work suggests that the nature of typed answers is different (certainly shorter) than written answers. The question naturally arises as to the adequacy of the typed answers for ranking purposes. Therefore, we shall perform a comparative study of the work of students who take both a mock electronic examination and the final written examination.

Summary
The feedback from the participants in the synchronous exam have confirmed our earlier findings that answering an exam at home can be a positive experience for many students. We have evidence that for students familiar with submitting assignments electronically, typing their answers at a computer need not be an impediment to completing an exam within the stipulated time. However, the rigorous time limit is a major issue for students in all forms of examination. It seems clear that anything that might result in a loss of time (slow typing, orientation, gaining access to the exam, periodically saving answers) is a major issue for all students. However, we were pleased to note that the overwhelming majority of students were positive in their reaction to electronic examinations as being “the way forward”, “a useful concept” and “a good idea”.

© British Educational Communications and Technology Agency, 2002.
The lengths of answers from all students were quite short—certainly shorter than would normally be expected in a conventional written exam. Even so, answers were sufficiently comprehensive for marking purposes.

We have identified the essential features that a synchronous, timed web-based exam paper must exhibit, including facilities for navigation, robustness and status but more can be done to improve the student experience through a redesign of the web-based paper.

References


