Chapter 7

Utilizing Automated Assessment for Large Student Cohorts

SYLVIA ALEXANDER, UNA O'REILLY, PAT SWEENEY and GERRY MCALLISTER
University of Ulster, Faculty of Informatics, Shore Road, Newtownabbey, Co. Antrim, N. Ireland BT37 0QB, UK. E-mail: u.oreilly@ulst.ac.uk

Assignment marking is a time-consuming activity, further exacerbated by the need for regular submission followed by informative and timely feedback. Online assessment methods have been adopted in order to address the pressing problem of coping with large student numbers. The system is used to provide a vehicle for the regular assessment of more than 400 students studying 'Networks and data communications.' This paper outlines the experience gained through large-scale use of automated assessment.

INTRODUCTION

Student numbers in higher education and in particular the demand for those with computing and information technology skills are increasing at a rapid rate which the supporting infrastructure within many computing science departments find difficulty in keeping pace with. Conventional support for academic lecturers in computing and information sciences is declining due to decreasing per capita student funding and a difficulty in attracting computing science research students, whose skills are generously rewarded in the commercial world. In particular, whilst the conventional classroom lecture can accommodate numbers limited only by physical space provision, the need for timely assessment and rapid feedback to large student groups presents a significant workload. The attractiveness of automated methods for testing and assessment increases accordingly.

To help meet the demand for skilled computing professionals many UK universities now offer non-computing graduates a postgraduate conversion course. This work relates to one such course, which has grown rapidly in popularity from 15 students at its inception in 1982, to a total registration of 572 in 2000/01, making it one of the largest courses of its kind in the UK. The course aims to provide education and training in the

concepts and methods of computing and information systems, relevant to the needs of the commercial, industrial and public sectors. The course is made up of 8 taught modules (6 compulsory and 2 optional) followed by a dissertation. It is offered in both full-time and part-time mode and taught on two campuses–Jordanstown campus, North of Belfast and Magee campus in Londonderry–with lecturers on both sites teaching to a common syllabus with identical assignments and examinations.

Habeshaw outlines a number of strategies for dealing with large student numbers [1]. Amongst the most attractive of these is the introduction of computer-assisted assessment. This paper outlines the experiences gained through the use of automated assessment in one module (Networks and Data Communications) of this course. The results quantify the effectiveness of this process, where CAA techniques were adopted to support the delivery of objective tests for both formative and summative assessment of students.

COMPUTER ASSISTED ASSESSMENT

As students become increasingly strategic in their study habits, there is much evidence to show [2] that assessment is the driving force behind student learning [3]. Students are motivated by feedback on their work, and regular formative feedback has been shown to have a marked improvement on students' overall performance [4]. To assist in improving learning, feedback needs to be timely and accurate. Increasing student numbers, coupled with a decline in resources, mean that staffs are unable to give formative feedback on student learning to the extent that they may wish. Understandably there is concern that as the quantity of marking increases, there is a corresponding deterioration in the quality of assessment. This has led universities to investigate new ways to assess students. Computer Assisted Assessment (CAA) offers the opportunity to assess students more regularly without increasing staff workload.

Computer Assisted Assessment encompasses the use of computers to deliver, mark and analyse assignments or examinations [5]. It also includes the collation of optically captured data gathered from machines such as optical mark readers.

CAA has many uses in Higher Education including [6]:

- Monitoring student progress;
- Testing a broad range of topics;
- Easing the marking load;
- Enabling more frequent testing;
- Early detection of individual and group problems and the corresponding identification of remediation;
- Detection of knowledge gaps, alerting academics of material not covered in sufficient depth.

In addition to the pedagogical advantages associated with using CAA techniques there are a number of important administrative benefits: [5]

- Marking is not prone to human error;
- Diagnostic reports and analysis can be easily generated

NETWORKS AND DATA COMMUNICATIONS MODULE

The module in ‘Networks and Data Communications’ provides an appreciation of the architecture, topology and protocols of local and wide area networks. In addition, it

provides an understanding of computer communication principles, an overview of the World Wide Web and applications such as e-commerce. Basic theory is augmented by information on the current technology that students may encounter in their workplace.

Currently students are taught in a conventional lecture environment, with assessment by both coursework and formal examination. Lectures were delivered over a 3-hour period each week, with 50% of the time being devoted to basic data communications and 50% to computer networks. The nature of the material allows for concurrent teaching of these topics without the need for pre-requisite knowledge of either.

In common with other modules on the course, registrations have risen sharply in recent years with a total of 413 students registered to take the networks and data communications module in 2000/01. Driven by the increase in student numbers, the module coordinators decided to incorporate an element of computer assisted assessment (CAA) into the coursework component of this module in order to limit the manual effort. The CAA tests were administered on-line using the 'Topclass' system [7], providing immediate feedback of results to both students and lecturers, and thus facilitating the identification of problems at an early stage.

EVOLUTION OF ASSESSMENT

Students are currently assessed using 3 methods:

1. A series of on-line class tests
2. A multiple-choice question paper (taken under formal examination conditions but automatically marked)
3. A conventional examination paper where students have to answer 2 questions (from 4) of a descriptive nature

Research has shown [8] that to be effective, the introduction and implementation of learning technology must be integrated within the structure and delivery of the course. Likewise, CAA must be implemented within an appropriate context. The purpose of the CAA must be apparent to the students; its aims and objectives should be clearly defined and the relationship between the CAA, teaching sessions and other assessed work clearly identified.

The following sections outline the techniques employed, providing commentary on how the CAA component has evolved to become fully integrated in the assessment process during the past three academic years.

Coursework

Authoring high-quality questions for objective testing is a time-consuming task and in its inaugural year (1998/99), only one CAA test was introduced. This test, covering approximately half of the syllabus, accounted for 15% of the total module mark. A further 10% was available via an individual written assignment, the remaining 75% being attributed to the examination. Despite the fact that only one test had been introduced, the time saved marking was considerable.

Based on the success of this system, in the following year the CAA component was increased and the allocation of marks redistributed such that coursework contributed 50% of the final mark for the module, the remaining 50% being scored in a 2-hour written examination.

Based on this model, in 1999/2000 the coursework was delivered as 5 class tests, each comprising of a number of questions based upon preceding lecture material. A final group work assignment, with students working in groups of 4, completed the coursework. The final coursework mark for the module was calculated as follows:

- The 4 best sets of marks from the on-line tests contributed 80% of the coursework mark
- The group work assignment contributed 20% of the final coursework mark (all members of the group receiving equal marks)

By discarding one set of marks from the tests, this ensured that any student who had not completed a test or had performed poorly (perhaps due to ill health) was not treated unfairly.

In the current academic year, it was decided that the group work component should be dropped, as part-time students found difficulty in cooperating in this type of activity. The on-line assessments continued to be delivered as 5 class tests, each comprised of a number of questions from preceding lecture material as demonstrated in Table 1.

Course-work #	Week Assessed	Material	Covered	Weeks Taught
		Networks	Data Communications	
1	3	Standards Encoding Systems	Asynchronous and Synchronous Communication Error Detection and Correction, Flow Control	1,2
2	5	Medium Access Control	Packet Switching, Datagrams and Virtual Circuits Series and Parallel communications	3,4
3	7	LAN Systems Ethernet, Token passing Wireless	Connections and Signaling Multiplexing	5,6
4	9	Network Interconnection	Networks Systems, Mediums and Topologies	7,8
5	11	All	All	9,10

TABLE I
COURSEWORK SCHEDULE

Examination

In 2000/01, the formal examination had 2 sections. Section A took the form of multiple-choice questions of a similar nature to the on-line tests, but was administered under strongly supervised examination arrangements. The students completed their answers using a 'survey form', which was then marked using Formic [9]. Section B was comprised of 4 descriptive examination questions, of which students were asked to complete 2 of their choice.

Students are required to submit handwritten evidence of their selection for any mathematically based question. This work is crosschecked with the completed survey form to ensure that no marks were allocated for randomly selected answers, thus ensuring that opportunist students did not benefit from inflated marks. Comparisons between

marks in Sections A of the examination and the TopClass tests are reported later in this paper in the 'Results' section.

TECHNOLOGIES

TopClass

WBTSystem's TopClass is a web-based solution for the delivery and management of online teaching and training materials. It operates as a client/server application over campus intranets or the Internet, providing administrative functions for managing learners, together with facilities for:

- Course delivery
- Discussion
- Testing

Registered students can access the system using any web browser in order to take courses and tests and engage in discussions with their peers and tutors.

The testing facility in TopClass creates tests from one or more question pools, where a pool acts as a repository for a set of questions. Each question must be associated with a question pool and pools can be used in multiple tests. Question pools may have just one single or multiple questions.

Local Implementation

The 406 students who enrolled to undertake the 'Networks and Data Communications' module were allocated to 'classes' within the TopClass system. A maximum of 100 students were able to access the system at any one time and students were assigned to sub-groups depending on their campus location and mode of study. Up to three laboratories, each containing 30 workstations were scheduled for sessions lasting 45 minutes each, in which the students accessed and completed the tests. A number of demonstrators familiar with the system helped in the supervision of the laboratories.

In week one of the module each student was provided with a username and a password with which they could access their own individual workspace throughout the term. A user-guide was also made available to the students to help them in navigating the system for the first time. Students accessed the system through a standard browser and entered their individual workspaces using the username and password provided.

The tests included a number of multiple-choice questions, ranging from the typical true/false Boolean question to list matching, where students had to carry out calculations and match their answer to one offered by the system. In an attempt to minimise students' attempts at plagiarism, pools of 4 questions, similar in style and level of difficulty, were created. One question from each pool would be displayed at random to each student.

After logging on students were allowed 35 minutes to complete and submit a test. If a student had failed to submit the coursework within the time period set the student was timed out of the system and no result was recorded.

Time restrictions were also placed on when students could access and submit tests. Typically tests were available for a 2-hour period. This was particularly beneficial for part-time students who may not have been able to attend lab classes due to work commitments. Such students could request permission to sit the test external to the

university. However, it was explained in advance to the students that the examiners hold the right to review the tests performed in unsupervised conditions with some scepticism.

Once the test window was closed, results for the auto corrected test were made available to the students the following day. Students were then able to access their results via the options in the menu presented following logon. After all the tests were completed they were made available to the students once again for revision purposes.

At any time students could use the discussion facility, either to communicate with each other or with the lecturer (although this facility was rarely used).

TopClass Limitations

The system implemented at Ulster has been utilising TopClass version 3.1. It is recognised that many of the problems reported in this section have been addressed in more recent software releases.

One of the major limitations with TopClass 3.1 is that it does not give students any advance warning as to when they will be timed out from a test. This is a major source of complaint from students as it results in a zero mark being recorded. In some circumstances students were allowed to re-sit the test.

Another factor was a limitation on numbers accessing the system concurrently. This necessitates sequential testing for a large cohort with associated logistical problems of invigilation.

Since the TopClass system was located on a server, which did not have 24/7 maintenance, it was inevitable that server problems would arise. On one occasion, this meant that the test had to be re-scheduled for the following week.

From an academic perspective, the greatest limitation of TopClass 3.1 is the type of results it returns. Whilst it is possible to obtain overall scores for the students it is not possible to see how the students score on individual questions. This presents a problem insofar as it is impossible to determine if there is a question that is presenting a problem to the class as a whole either because

- The incorrect answer has been flagged as the correct answer or
- The class has had difficulties in interpreting the lecture materials and requires remedial assistance from the lecturer.

The most recent release of TopClass (version 5) addresses many of these issues, providing comprehensive tracking facilities and reports on class and individual learner progress. Teachers can easily monitor the progress of their assigned learners individually or by class. TopClass also integrates with Oracle Reports for powerful reporting capabilities.

Formic

Formic is an automated data entry and validation forms processing system. It automatically reads and processes any combination of tick boxes and handwriting offering a complete data collection solution. An integrated statistics module provides facilities for statistical analysis and enables the user to view the survey results database in various graphical formats. The system also provides for export of comma-delimited files to other software packages.

STUDENT INTRODUCTION

In order to acquaint students with the features of the TopClass system and the different types of questions available to them, their first exposure to the system was an on-line (non-evaluated) survey, to determine gender, age, qualifications, previous educational institution, previous computing experience, ethnic origin and any disability factors. Whilst this was not relevant to the outcome nor considered in the performance measurement, it provided valuable profile information for other uses in overall module evaluation and allowed a convenient mechanism for introduction to the system.

RESULTS

On comparison of Section A of the examination with the on-line assessment scores achieved by students on this module the following observations can be made.

Students on average scored 14 fewer marks in Section A of the examination compared to their on-line assessments. The distribution of each examination section and assessment scores is shown in Figure 1.

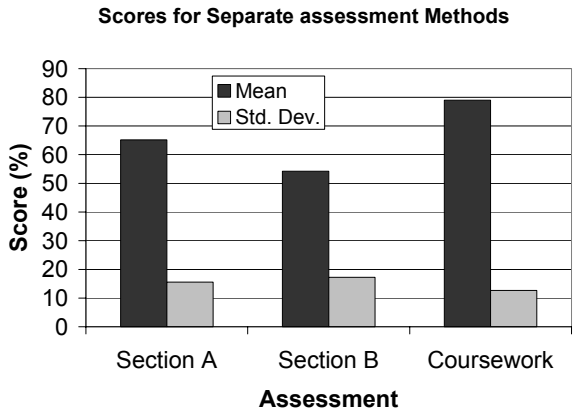

FIGURE 1

Some 5% of students achieved an examination mark that was more than 2 standard deviations less than the mark that they had achieved in the coursework. There is obviously concern that these students may have achieved an inflated score in the coursework through guessing.

However, there was a similar proportion of students for whom the situation was reversed, i.e., they scored a mark more than two standard deviations higher in the examination than the coursework. It would appear that these students progressed considerably despite the fact that they had not adapted well to the continuous learning that was encouraged throughout the delivery of the module.

The difference of 14% between the average score in assessment and examination is similar to that on other modules of the course and is therefore not a cause for overdue concern. The following factors contribute to this effect:

- The equal weighting of each test, earlier tests are easier, for encouragement;

- Students have less information to retain for a shorter time in the test scenario;
- Assessment is continuous and drives students to score well;
- By its nature the assessment process is not as secure as the examination process.

Section B of the examination used traditional examination questions giving students a choice of 2 questions from 4. In each case the question was comprised of sub-sections, which got progressively more difficult so that differentiation of the students' abilities could be determined. Overall there was an average difference of 24% between coursework and this method of examination.

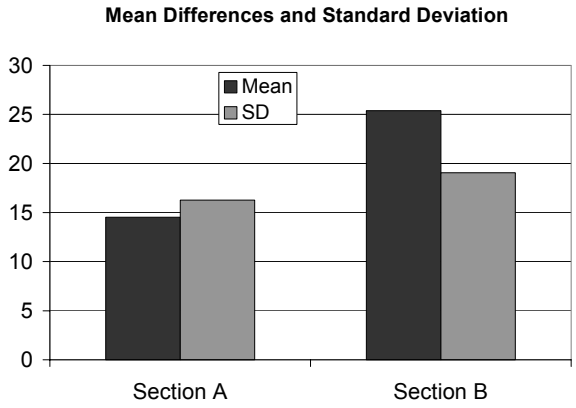

FIGURE 2

Figure 2 shows the mean differences between each examination section and coursework. Although overall scores are comparable there is a large standard deviation - particularly in Section A, indicating that issues other than ability may be influencing the results.

The coursework, although used in a summative manner, was mainly formative in nature and most students performed well, achieving a high average. Section B of the examination produced a more traditional distribution with a wider spread of results.

Section A and Section B of the examination measures different skills possessed by students. Section A (the multiple-choice section) requires the student to provide a response to a question where the answer is pre-determined - testing specific skills and levels of competence.

The more traditional examination questions presented in section B require students to combine knowledge, understanding, and skills, to solve a problem or present an argument. Students generally scored higher in Section B. In a similar study, Kumar [10] noted a strong correlation between performance in written examination and on-line test scores. Similarly, students who performed well in the coursework also performed well in the examination and in the other modules of this course.

Student Perceptions

Surveys carried out by Griffiths et. al. [11] show conclusively that students prefer computer-based tests to paper-based methods. Students undertaking the module in Networks and Data Communications shared this view. Despite the fact that few students had previous exposure to CAA (Table 2), most found that this approach provided a useful learning experience.

	CAA (%)	TopClass (%)
Yes	16	3
No	84	97

TABLE 2
PREVIOUS EXPOSURE TO CAA

The majority found the system acceptable to use despite some initial problems encountered in the access and submission processes (Table 3).

	Navigation (%)	Access & Completion (%)
Difficult	3	5
Satisfactory	37	54
Easy	33	33
Very Easy	27	8

TABLE 3
USE OF TOPCLASS

Of the total student cohort 35% felt that 5 tests was too many, the remaining 65% felt that the frequency was about right. Unsurprisingly, no one felt the need for more frequent testing. Almost everyone (95%) was in favour of using the 4 best sets of marks. 27% of students found the TopClass tests difficult, the remainder felt they were pitched at the right level and none felt that it was an easy option.

In terms of the student experience the majority felt that it had both improved the learning experience and assisted in their time management, with most students requesting that this approach be adopted in other modules (Table 4).

	Improved learning experience (%)	Improved time management (%)
Yes	78	81
No	8	5
No difference	14	14

TABLE 4
STUDENT LEARNING

The overall feeling was that this type of assessment was at least satisfactory in producing a realistic and reliable grade and in assisting the understanding of the syllabus content (Table 5).

	Realistic and Reliable Grading	Aid to Understanding
Poor	11	8
Satisfactory	38	32
Good	40	41
Very good	11	19

TABLE 5
STUDENT EXPERIENCE

Individual responses from students indicate that this method of assessment has other attractions. One student indicated that it was more 'time efficient than assignments' whilst another expressed the opinion that it helped to 'focus study throughout the module.' Most criticisms related to the "element of luck" involved in MCQs, which could give students a "false sense of confidence." The problems relating to tests being timed out and the resulting zero mark were also a cause for concern.

CONCLUSIONS

From an academic point of view this approach would appear to be reasonably successful, with the majority of students achieving good scores in both coursework and examination. However, one must recognise a 14% and 24% discrepancy with Sections A and B respectively of the formal examination. Listed below are some of the issues which the course team believe can reduce this for the future:

1. The effect of guessing is a cause for concern. This can be largely eliminated through the use of corrective scoring in which marks are subtracted for incorrect answers–this option may be considered.
2. The coursework could be made more demanding by the introduction of alternative question types e.g. permutational multiple-choice questions [12], [13]. More powerful software will enhance the ability to provide both more challenging questions and additions to the pools for each one.
3. All three methods of assessment have legitimacy since they measure different quantities. Nevertheless, further investigation regarding the weighing of each is required.

At a logistical level, increasing laboratory space will eliminate the need for sequential testing, allowing more thorough invigilation of the on-line activity.

Overall, evidence based on these results indicates this as an acceptable means of assessment, when compared to the usual manual process. The results must be considered in light of the fact that in general coursework attracts higher marks than those obtained from a formal examination setting. However, in summary, greater challenges must be set in both the on-line tests and in the Section A examination questions.

The way forward

The future of CAA lies in reinventing assessment such that CAA is integrated seamlessly with other teaching methods [14]

From 2001/02 WebCT [15] will be adopted for the delivery of on-line assessment in this module. In addition to addressing some of the identified deficiencies of the current system, an institutional strategy has been adopted to streamline and integrate administrative functions with a valid and reliable assessment system. Adoption of the new system will provide opportunities to diversify and enrich the manner in which the assessment is managed and delivered. For example, WebCT facilitates the provision of calculated questions which enables the generation of mathematical questions based on a random set of variables. Once the question and the formula are entered, a series of questions (up to a maximum of 100) are automatically generated using the chosen parameters. Students will therefore receive a wide variety of different questions based on the single application of a formula thus increasing the security of the exercise.

ACKNOWLEDGEMENTS

The authors would like to thank the staff of the Educational Development Unit of the University of Ulster for the use of their server and in particular Dr. Alan Masson for his expert advice on the use of the TopClass system.

Thanks are also due to the Staff Development Unit for the training in and use of Formic.

REFERENCES

1. Habeshaw S. et al., "Problems with Large Classes: Making the Best of a Bad Job." Technical and Education services. Bristol (1992).
2. Kneale, P. and Collins C. "Study Skills for Psychology," London, Arnold (1996).
3. Brown, G, Bull, J and Pendlebury M, *Assessing Student Learning in Higher Education* Routledge, (1997).
4. Schmidt N, Norman, G and Boshuzen H, "A Cognitive Perspective on Medical Expertise: Theory and Implications," *Academic medicine*, 65, (1990).
5. McKenna, C and Bull, J, "Frequently Asked Questions about CAA," CAA Centre http://www.caacentre.ac.uk/faqs.shtml (2001).
6. Stephens, D and Mascia, J, "Results of a Survey into the Use of Computer Assisted Assessment in Institutions of Higher Education in the UK." (1997).
7. WBTSystem's TopClass, http://www.wbtsystems.com/products/products.html
8. Ehrmann, S C, 1998 "Studying Teaching, Learning and Technology: a Toolkit from the Flashlight Program," *Active Learning* 9, (1998).
9. Formic, http://www.formic.co.uk/
10. Kumar, A., "On Changing from Written to On-line Tests in Computer Science I: An Assessment," SIGCSE Bulletin – Inroads, *ITiCSE'99 Proceedings*, Vol 31, Number 3 (1999)
11. Griffiths, F., "Freedom and Flexibility in the Higher Education Examination Process," *Active Learning* 1 (1994).

12. Farthing, D., Jones, D. and McPhee, D., "Permutational Multiple-Choice Questions: An objective and Efficient Alternative to Essay-Type Examination Questions." *ITiCSE, 98 Proceedings*, (1998)

13. Bush, M., "Alternative Marking Schemes for On-Line Multiple Choice Tests," 7th Annual Conference on the Teaching of Computing, (1999).

14. Bennett, R.E. "Reinventing Assessment: Speculations on the Future of Large Scale Educational Testing," Princeton, NJ: Information Centre, Educational Testing Service, (1998).

15. WebCT, http://www.webct.com

Sylvia Alexander received the B.Sc. (Hons) degree in Computing and Applied Mathematics from Queen's University, Belfast, N. Ireland in 1987. She received the M.Sc. degree in Informatics from the University of Ulster in 1994. Following a period in industry as a Systems programmer she moved into the field of Educational Technology where she has remained for the past 13 years. In 1990 she joined the Computers in Teaching Initiative Centre for Computing based at the University of Ulster as the Centre Manager and in 2000 secured funding for the Learning and Teaching Support Network Centre for Information and Computer Sciences. In 2000 she was appointed as a Lecturer in Informatics. Her main areas of interests are in the area of computer science education research, particularly pedagogic and technological innovation.

Una O'Reilly received the B.Sc. (Hons) degree in Information Management from Queen's University Belfast, N. Ireland in 1997 and received the M.Sc. degree in Computing and Information Systems from the University of Ulster in 2001. In 1998, she joined the national CTI Centre for Computing based at the University of Ulster and is currently an Information Officer with the Learning and Teaching Support Network Centre for Information and Computer Sciences. Her main interests are in the field of Computer Assisted Assessment and Technological Innovation in Computer Science Education.

Patrick J. Sweeney received the B.Sc. (Hons) degree in Physics and Chemistry from the Ulster Polytechnic, Belfast, N. Ireland in 1975. He received the M.Sc. in Opto-Electronics from Queen's University, Belfast in 1976. He received the Post Graduate Certificate in University Teaching in 1999 from the University of Ulster, N. Ireland. In 2000 he joined the Institute of Learning and Teaching. He holds the post of Director of Teaching and Learning in the Faculty of Informatics. He spent four year in research on the Optical measurement of the Neutral Winds of the Polar Thermosphere. In 1980 he entered his current employment as a Lecturer in Computer Science and in 1990 became a Senior Lecturer. His main interests are in Data Communications and Computer Architecture with an emphasis upon technology transfer to indigenous industry. He has publications ranging though atmospheric physics, parallel processing and educational issues.

Gerry McAllister obtained a degree in Electrical and Electronic Engineering in 1977, a M.Sc. in Electronics in 1983 and a Ph.D. in 1996. After graduating he was employed as Senior Research Officer in Queen's University, Faculty of Medicine. During this time his main interests were in recording methods of EEG and Evoked Potentials. He is a Chartered Engineer and a Member of the Institution of Electrical Engineers. In 1987 he was employed as Lecturer in Informatics at the University of Ulster and appointed Senior Lecturer in 1993. His current main research interests are in new methods of detection and correction for hearing acuity and the use of technology in teaching and assessment. He has over 55 publications in conferences and journals, within the remit of Bio-Medical Engineering and Teaching and Assessment. He is the Director of the National LTSN centre for Information and Computer Science.