

**STRATEGIES FOR LARGE-SCALE  
ASSESSMENT:  
AN INSTITUTIONAL ANALYSIS OF  
RESEARCH  
AND PRACTICE IN A VIRTUAL  
UNIVERSITY**

**G Sim, N Malik and P Holifield**

# Strategies for Large-Scale Assessment: an Institutional Analysis of Research and Practice in a Virtual University

Gavin Sim<sup>a</sup>  
Department of  
Computing  
University of  
Central Lancashire  
Preston  
PR1 2HE  
Tel: 01772 893275  
Email:  
grsim@uclan.ac.uk

Dr Naveed Malik<sup>b</sup>  
Virtual University  
3rd Floor, Building #1-2  
Aiwan-e-Iqbal complex,  
Egerton Road  
Lahore.  
Pakistan.  
Email:  
rector@vu.edu.pk

Dr Phil Holifield<sup>c</sup>  
Faculty of Design  
and Technology  
University of Central  
Lancashire  
Preston  
PR1 2HE  
Tel: 01722 893168  
Email:  
pholifield@uclan.ac.uk

<sup>a</sup>Department of Computing  
<sup>c</sup>Faculty of Design and Technology  
University of Central Lancashire  
Preston  
PR1 2HE

<sup>b</sup>Virtual University  
Lahore  
Pakistan

## Abstract

The Virtual University in Pakistan is a major government initiative to provide higher education to increasing numbers of students who are qualified but cannot afford fees or find places at existing conventional universities. Administering large-scale assessment is problematic for distance learning courses with issues such as authenticating the users and the increase burden of marking. These problems are heightened with the dispersion of the student cohort and the adoption of computers to the assessment process.

The Pakistan Virtual University have devised an assessment strategy that enables assessment to be administered with both a large and dispersed student cohort. This paper aims to examine the policies and strategies adopted by the VU and the infrastructure required to fulfil these. The infrastructure is examined at national and institutional level as both have a direct impact on the strategy. The infrastructure and strategy will be analysed against research into the adoption of Computer-Assisted Assessment. Conclusions are drawn as to the success of the current strategy and future likely developments.

**Keywords:** large-scale assessment, strategy, infrastructure

## **Introduction**

The Virtual University (VU) in Pakistan aims to exploit the potential of Information Technology by utilizing the services of leading academics within and beyond the geographical boundaries of Pakistan to produce and deliver courses. Each academic records a series of lectures, and authors assignments, exams and other materials. Through use of satellite television and a web-based Learning Management System (LMS) courses are offered to students at learning centres, which they attend daily. There is evidence that the needs of learners in remote areas can be served via local learning centres supported by web-based learning (Mayes, 2001). However based on this model and with limited academic staff available, administering large-scale assessment is a challenge.

The current first semester cohort is in excess of 1700 students and future expansion is expected to generate an annual cohort of 30,750 by 2005 (Knight, Malik et al. 2000). This has resulted in the assessment strategy adopted by the VU being a significant departure from the assessment model of Bachelors and Masters programs conducted by conventional universities in Pakistan. In such programmes, assignments carried out by students do not form part of the assessment for the course, rather a full year's lectures are followed by comprehensive examinations. With the large cohorts of the VU, this assessment strategy is not viable due to limited academic staff. For example, a typical computer essay exam scripts may take half an hour to mark, for 1700 students taking both a mid-term and final exam, a total marking time of 1700 hours (212 days) is required for a 1 semester module. Also because of the distance learning nature of the VU course, and the lack of face-to-face contact with expert tutors, a different assessment strategy has been designed.

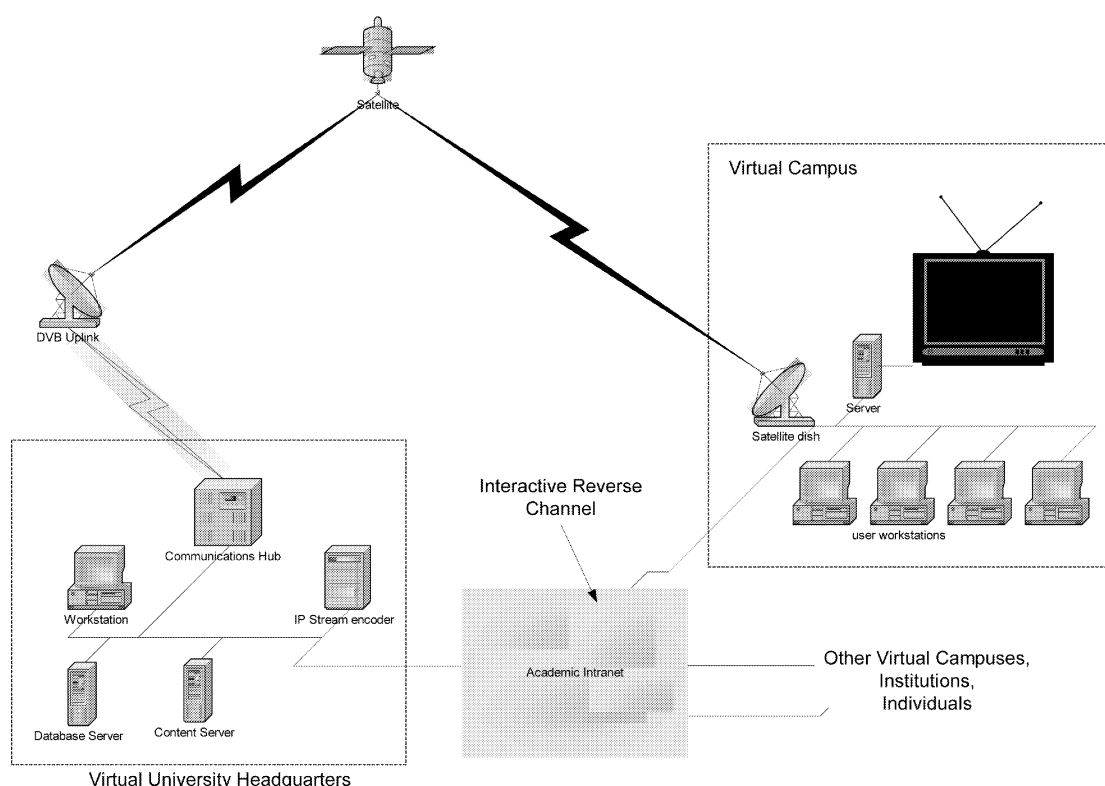
The VU has established a policy of summative assessment through assignments, mid-term and final exams. The policy is based on assessment strategies used by traditional American Universities with the addition of computer-assisted assessment. This policy enables students to monitor their own progress, helps maintain motivation, and promotes active learning. The assessment strategies and policies are crucial factors that not only influence students' performance, but determine the credibility of the VU qualifications within Pakistan and internationally.

## **Infrastructure**

To enable the successful implementation of both a CAL and CAA strategy the infrastructure is a crucial component. An understanding of the technical requirements for delivering the courses and assessment is necessary to prevent potential difficulties. Risk assessment and management are vital in developing an assessment strategy encompassing CAA (Zakrzewski and Bull 2002). Recommended guidelines for an institutional strategy for CAA highlight the need to identify technical issues (Stephens, Bull et al. 1998). This focuses mainly on the analysis of software but with a distributed student cohort other technical features such as connectivity need to be examined. Such features are affected by national IT infrastructure and cannot be solved by a single institution. The issues surrounding bandwidth to certain locations across

Pakistan are being addressed through the government's policy to develop an Academic Intranet.

In order to administer courses and assessments across various geographical locations the use of an LMS system is required. The LMS adopted by the VU is a commercial product from Intralearn that is also being used by the University of Massachusetts. The technology and infrastructure the VU have adopted to administer the courses are represented in Figure 1.



**Figure 1. The infrastructure used by the VU to deliver courses.**

Intralearn requires a ColdFusion back end and the user interface is entirely browser-based. The backend requires a DBMS for which Microsoft SQL-Server is used. Intralearn and SQL-Server are hosted on separate servers creating a 3-Tier system. This is necessary due to the high volume of traffic to the site and enables a distribution of processing resources. This approach also allows increased security of the student records database.

Students interact with the LMS over the Internet, for which purposes the VU have obtained 2Mbps bandwidth between local centres and the VU hub. So far a peak usage of 1.9Mbps has been observed. An important consideration is the adequacy of the bandwidth available to meet future expansion because if this cannot be increased, growth in student numbers becomes unfeasible. To enable scalability, the VU can increase the bandwidth available on demand as the connection is via an Ethernet port of the router that is on a 155Mbps connection straight to New York.

The software used to deliver courses must be subject to pedagogical critique (Hawkrige, 1993). Expanding on this, software engineering criterion may be used to evaluate the systems through ISO standards (Valenti, Cucchiarelli et al. 2002), BS standards (Kleeman and Osborne 2002), and user requirements (Sclater and Howie 2003). One of the important considerations of the VU was the functionality and reliability of its software as experienced by its students. Although it is possible to integrate commercial CAA applications into its LMS, the VU developed in-house software to deliver its assessment strategy.

As students are not centrally located assignments need to be distributed electronically. Assignments are posted and submitted through a custom built application integrated into the LMS. This method avoids undertaking critical activities at specific times because of the variability of Internet connectivity across the country, and allows students to work asynchronously preventing data loss due to technical issues.

Due to the nature of some assignments text needs to be formatted in a particular manner for example, maths symbols. The application enables the students to use whatever tool is available and then upload the file. On submission it is encrypted, renamed by the application and presented for grading to the tutor. This helps to alleviate potential security flaws of sending emails across the Internet such as authentication of user and encryption (Hatton, Boyle et al. 2002).

The delivery and submission of exam papers via the Internet to the learning centres is an issue that has security implications. The exam papers need to be stored on a secure server and students have access to the paper only for the duration of the exam. With such a large cohort conducting the exam synchronously via the LMS would place considerable pressure on the servers therefore an asynchronous approach is required. Custom software that integrates into the LMS has been developed to enable the delivery of exams in a secure environment. This software is deployed at designated exam centres across Pakistan. The question paper is sent as an encrypted file via e-mail on the day of the exam to the centres. The examiner puts the file in a special folder on the server at each site and conducts the exam. The same file is returned immediately after the exam via e-mail, again encrypted for security reasons.

The technical infrastructure and software considerations required to administer large-scale assessment strategies have been examined. These issues need to be evaluated in conjunction with the development of the institutional strategy. Any change in assessment procedures could have technical implications that need to be addressed.

## **Assessment Strategies**

The admissions exams are the first contact the students have with the assessment methodology adopted by the VU. The admissions are scheduled for spring and autumn enabling two cohorts of students a year as opposed to the single enrolment within UK universities. The admissions test was administered nationwide for the autumn 2002 admissions. This was a paper-based exam with a mixture of MCQ and essay style questions. Due to the number of students involved (5000) the marking of these scripts was outsourced to the National Testing Service and the results communicated electronically to the VU. Although the results were good, no significant advantage was gained by using the test as an admission criterion. One primary cause for concern was that whereas students from the larger towns and cities were at least exposed to admission tests, it was a brand new experience for students from the smaller towns. This has a built-in bias towards the already empowered. As a result the VU decided that subsequent admissions would be based solely on the marks obtained in the Matriculation and Intermediate public examinations, thereby providing a level playing field to all students regardless of their geographic location.

The number of assignments varies between modules but is usually 5-10 per semester module. The assignments account for 15% of the overall student grades for the module. Even though they are small percentages for each assignment it relieves some of the inevitable pressure that occurs when the assessment is reliant on high stake exams as used traditionally in Pakistan. In an open learning environment regular testing is vital (Croft, Danson et al. 2001), however, a series of bad marks can lead to a downward spiral for students learning and confidence (Taras, 2002). Therefore it is necessary to build the students confidence from an early stage in the course and the assignment strategy adopted helps achieve this. The first assignment for the CS101 module, Introduction to Computing, simply requires the students to send an email from their VU account to a specified account. Although this may appear to be a straightforward activity, it was designed to ensure the students are familiar with the LMS and build confidence from the outset. As research suggests that students with more computer experience have lower anxiety and better attitudes towards using technology within exams (Liu, Papathanasiou et al. 2001). This factor needs to be considered as test anxiety can have an adverse effect on students' ability to perform in exams and significantly lower their grades (Cassady and Johnson 2002). Subsequent assignments relate to the learning outcomes and syllabus of the modules. The academics who design the modules syllabus are also responsible for the assignment and exams.

As previously explained, assignments are distributed and submitted via the LMS. Objective testing is not usually used for the assignments, an essay style format is mainly used. This places marking pressure on the tutors requiring continuous diligence, and the overall management has to be kept tight to ensure no delays. Typically one tutor can grade assignments from 250-300 students per week. This is possible because there is no face-to-face interaction with the students or classroom activities. The number of tutors marking the assignment is proportional to the number of students, for

example, currently with the English module there are 5 tutors and 1 Professor all employed on a full time basis. The tutors can be expanded in conjunction with the student numbers to maintain this policy as no additional professors need be recruited.

Typically, solutions to the assignments are posted on the LMS the day after the due date. Thus, late assignments are not accepted. Often messages are posted on the discussion board concerning late submission of assignments a number of claims are due to technical faults. The VU policy enables an extra day after the hand-in date to take account of any technical difficulties experienced at the Learning centres. After this date no assignments are accepted.

So far, no group or individual feedback has been done on the basis of assignments. Students are required to monitor their own progression as part of an active learning strategy. Through self-reflection on their performance against the model answer provided students can gain a personal insight into their understanding of the subject (Peat, Franklin et al. 2001). However, students have the opportunity to ask questions to the tutor concerning the assignment via the moderated discussion board (MDB) and all students will have access to the response. Although this does result in some individuals obtaining feedback no unfair benefit is expected as everyone has access to the comments. Therefore, a group policy could easily be adopted for feedback via the MDB and this may further enhance the student experience, as feedback is an effective means of influencing learning results in a positive way (Gordijn and Nijhof 2002).

There are two exams per module, a mid-term and final exam with weightings of 35 and 50 percent. These exams are conducted at designated examination centres by examiners appointed by the VU who are responsible for ensuring student identity and authentication.

The actual exam is divided into two sections the first being a MCQ section and the remaining section is composed of essay style questions. The first section is automatically marked via the computer. The MCQ are presented in the browser as series questions with radio buttons for the options enabling automatic collection of responses and grading. This helps alleviate some of the burden placed on the staff through marking exam scripts. However, the second half of the paper is an essay style question and human intervention is required to mark the paper. The fact that student responses are typed has the benefit of legibility and comprehension for the marker (Thomas, Price et al. 2002), therefore time is not spent deciphering students handwriting making the process more streamlined. The strategy of using both MCQ and essay-style questions has the advantage of being able to test a range of cognitive skills as outlined in 'The Taxonomy of Educational Objective' (Bloom, 1956). Questions concerning the viability of testing the higher cognitive skills of synthesis and evaluation through objective testing (Sim, Holifield et al. 2003) can be addressed by a two-section exam. The essay format helps address this issue and other problems associated with MCQ such as guessing (Harper, 2003).

The results of the exams are announced through the LMS and again no individual feedback is provided to the students. Should students have concerns about their marks they can contact the VU staff who will try to resolve their problem.

With a large-cohort of students scalability of this format of exam is a concern especially for students entering the final year. Emphasis is placed on the assessment of higher cognitive skills and it is debateable whether objective testing is feasible across all subject domains. Paterson indicated that it is not feasible to test the higher-level cognitive skills within mathematics (Paterson, 2002). Bloom states that in the majority of instances Synthesis and Evaluation promote divergent thinking and answers cannot be determined in advance (Bloom, Hastings et al 1971). However, Reid (Reid, 2002) and Daly & Waldron (Daly and Waldron 2002) conclude that the higher-level skills can be assessed via CAA. With the advancement in technology surrounding automated essay markers it may in future be possible to examine large student cohorts producing essay responses with computer aided assessment.

Projects are being planned for the final year. A concept is being developed that requires students to work in cyber-space resulting in no two-team members belonging in the same city. This in itself should provide valuable experience in the modern work environment. The same staff will grade these assignments but over a longer period of time. Explicit guidelines for grading these projects will be developed to simplify the task. The only gap that remains is the face-to-face presentations or lab demonstrations no solution has currently been found that scales with large cohorts.

## **Conclusions**

The development of a system to administer large-scale assessment to a distributed student cohort at learning centres across Pakistan has been examined. This has highlighted the need for a strategy and policies to consider the technological resource implications and infrastructure necessary to implement these systems. A great deal of resource has been required to implement the VU's strategy from servers to an Academic Intranet. These factors influence not only the decision making process of the VU, but the same considerations are important even at departmental levels within universities adopting CAA.

The strategy adopted by the VU has shown that a mixture of CAA and tutor centred marking is achievable with the commitment of dedicated tutors. This does have its limitations with the provision of individual feedback but this has not proved to be an issue. Currently no viable methods of integrating lab demonstrations or presentations for large cohorts of students have been identified.

Due to the fact that the VU has been operational since March 2001 the investigation can only draw preliminary conclusions. There are no students in the third and fourth year of the course so investigations concerning the viability of the strategy to test achievement at the latter stages of an honours degree cannot be made. In addition, the way that this model of assessment



functions when students can enrol and participate from home needs to be examined. However, early analysis has shown how large-scale assessment can be implemented successfully within Pakistan through the use of a LMS and learning centres.

## **References:**

- Bloom, B. S. (1956) *Taxonomy of Educational Objectives: The classification of educational goals. Handbook 1. Cognitive Domain* Longman.
- Bloom, B. S., Hastings, J. T. & Madaus, G.F. (1971) *Handbook on Formative and Summative Evaluation of Student Learning* McGraw-Hill Book.
- Cassady, J. C. & Johnson R. E. (2002) *Cognitive Test Anxiety and Academic Performance*. Contemporary Educational Psychology **27**(2) 270-295.
- Croft, A. C., Danson, M., Dawson, B.R. & Ward, J.P. (2001) *Experience of using computer assisted assessment in engineering mathematics*. Computers & Education **37**(1) 53-66.
- Daly, C. & Waldron, J. (2002) *Introductory programming, problem solving and computer assisted assessment*. Proceedings of the 6th International Computer Assisted Assessment Conference, Loughborough.
- Gordijn, J. & Nijhof, W. J. (2002) *Effects of complex feedback on computer-assisted modular instruction*. Computers & Education **39**(2) 183-200.
- Harper, R. (2003) *Correcting computer-based assessment for guessing*. Journal of Computer Assisted Learning **19**(1) 2-8.
- Hatton, S., Boyle, A., Byrne, S. & Wooff, C. (2002) *The use of PGP to provide secure email delivery of CAA results*. Proceedings of the 6th International Computer Assisted Assessment Conference, Loughborough.
- Hawkridge, D. (1993) *Challenging Educational Technology* Athone Press.
- Kleeman, J. & Osborne, C. (2002) *A practical look at delivering assessment to BS7988 recommendations*. Proceedings of the 6th International Computer Assisted Assessment Conference, Loughborough.
- Knight, P. T., Malik, N. A. & Iftikhar, A. (2000). *Feasibility Study for the Pakistan Virtual Information Technology University and the South Institute of Information Technology* <<http://www.knight-moore.com/projects/SIIT-VITUreport1.html>> (28<sup>th</sup> June 2002)

- Liu, M., Papathanasiou, E. & Hao, Y. (2001) *Exploring the use of multimedia examination formats in undergraduate teaching: results from the fielding testing*. Computers in Human Behaviour **17**(3) 225-248.
- Mayes, J. T. (2001) *Quality in an e-University*. Assessment & Evaluation in Higher Education **26** (5) 465-473.
- Paterson, J. S. (2002) *Linking on-line assessment in mathematics to cognitive skills*. Proceedings of the 6th International Computer Assisted Assessment Conference, Loughborough.
- Peat, M., Franklin, S. & Lewis, A. (2001) *A review of the use of online self-assessment modules to enhance student learning outcomes: are they worth the effort of production?* Proceedings of ASCILITE 2001, Melbourne.
- Reid, N. (2002) *Designing online quiz questions to assess a range of cognitive skills*. Proceedings of the World Conference on Educational Multimedia, Hypermedia and Telecommunications, Denver.
- Sclater, N. & Howie, K. (2003) *User requirements of the 'ultimate' online assessment engine*. Computers & Education **40** (3) 285-306.
- Sim, G., Holifield, P. & McManus, B. (2003) *Online assessment of multimedia: from Knowledge to Evaluation the problems of classification*. Proceedings of CAL03, Belfast.
- Stephens, D., Bull, J. & Wade, W. (1998) *Computer-assisted Assessment: suggested guidelines for an institutional strategy*. Assessment & Evaluation in Higher Education **23** (3) 283-294.
- Taras, M. (2002) *Using Assessment for Learning and Learning from Assessment*. Assessment & Evaluation in Higher Education **27** (6) 501-510.
- Thomas, P., Price, B., Paine, C. & Richards, M. (2002) *Remote electronic examinations: student experience*. British Journal of Educational Technology **33** (5) 537-549.
- Valenti, S., Cucchiarelli, A. & Panti, M. (2002) *Computer Based Assessment Systems Evaluation via the ISO90126 Quality Model*. Journal of Information Technology Education **1** (3) 157-175.
- Zakrzewski, S. & Bull, J. (2002). *The mass Implementation and Evaluation of Computer-Based Assessments*. Assessment & Evaluation in Higher Education **23** (2) 141-152.