# Production and delivery of TRIADS Assessments on a university-wide basis

Don Mackenzie Centre for Interactive Assessment Development University of Derby Kedleston Road Derby, DE22 1GB

## Abstract

The Centre for Interactive Assessment Development was set up at the start of the current academic year to produce and deliver interactive assessments across the University of Derby as a whole. This is the first time that the TRIADS system has been used on such a large scale.

The experience gained in establishing operational procedures, protocols and working relationships during the year are discussed and examples of question styles delivered in a variety of disciplines are demonstrated.

## Introduction

The University of Derby formally recognised over ten years of experience in the production and delivery of computer-based assessments in the School of Environmental and Applied Sciences by creating a Centre for Interactive Assessment Development (CIAD) in November 1999. During the first year of operation the Centre had a staff of 3 full time equivalent members covering both academic and technical backgrounds.

The Centre provides a service for the design, production and delivery of computer-based assessments across the university and for industrial clients. The Tripartite Interactive Assessment Delivery System (TRIADS) is exclusively used for production and the Centre additionally provides TRIADS system support and development for the thirty-seven TRIADS evaluation departments covering eighteen disciplines in twenty-five universities. (HEFCE-FDTL funded project in collaboration with the University of Liverpool and the Open University.)

This paper highlights some of the issues relating to the development and dissemination of CBA beyond the 'cottage industry' stage.

## Dissemination

During the first seven months of operation the Centre will have produced and delivered over 2500 student assessments in the areas of Biology, Earth Science, Geography, Environmental Science, Electrical Engineering, Business Economics, Computing, Radiography and Modern Languages.

Initial production centred on those areas of Environmental and Applied Sciences that had previously been delivering an increasing number of assessments. However a high profile opening of the Centre by the Chancellor of the University coupled with open days and presentations in some Academic Schools led to a widening of interest and persuaded a number of academic staff to try CBA for the first time with favourable results. The Centre is not yet working to full capacity when viewed overall but dissemination activity this year was deliberately less intensive in order to facilitate commissioning of equipment and development of the assessment system itself.

More intensive dissemination is planned to start after the examination period in an attempt to provide an even work-flow through next semester.

#### **Production issues**

Each assessment was allocated to an individual member of the production team on receipt, completed in consultation with the tutor and independently checked by a second member of the team prior to a final check by the tutor.

All assessments were produced and delivered on time, however a number of production issues arose during the period. These were experienced and overcome this year but will need more focussed attention next year. The most significant of these are discussed below.

#### Scheduling and bunching of assessments

Course delivery at University of Derby in common with many institutions is organised on the basis of a two-semester year with examinations for the first semester scheduled for the last three weeks of January. Virtually all modules run for a single semester and thus computer-based assessments associated with coursework exercises tend to be concentrated towards the end of each semester unless they run during the examination period. These factors mean that there is a danger that assessments might not be produced unless questions are supplied in sufficient time for them to be input into the system. Furthermore, such concentration increases the temptation for tutors to use or minimally update a previous assessment.

Despite our best intentions to generate a balanced work-flow, academic staff are now under significant pressures and many questions and even whole assessments have been received at the last minute. This year we have been able to be flexible in our approach but a more formal cut-off deadline for receipt of assessments will have to be implemented if the predicted increase in number of assessments can be completed successfully next year.

#### Assessment review & multiple contributors

Even if assessments materials are received in time to facilitate production, last minute changes by question authors can create scheduling problems. To some extent this is a function of the inexperience of tutors with this mode of assessment delivery. A question can look very different on screen compared to the traditional paper version and ambiguities can easily be overlooked at the planning stage. Formal review deadlines are thus absolutely necessary. The problems of keeping to review deadlines are compounded where a single assessment has multiple contributors.

## Presentation of questions for input

Standard question types in TRIADS are accompanied by Question Definition Proformas which, when completed, contain unambiguous information about the question content, mode of operation, scoring and feedback. Question Definition Proformas may be completed in any word-processing package and when saved in ASCII format can be copy/pasted directly into the TRIADS system.

Whilst some tutors have used these successfully, some have completed them inappropriately, some are completely baffled by them and others have presented questions in a variety of alternative formats.

Developments in the TRIADS system will allow some question styles to be completed interactively by tutors, however the more complex question styles supported by the system are difficult for tutors to generate in this manner and a more structured method of question submission is required.

#### Academic staff development issues

Since CBA is a new assessment method for many, it has become apparent that a significant programme of staff development is required in the area of assessment and question design for electronic delivery. This is particularly so with respect to the definition of learning outcomes being assessed and the balance of questions in a test with respect to facility, scoring and content. TRIADS provides the functionality to finely tune the scoring in many of its question types and it is important that tutors understand the implications of their decisions on the final mark distribution. Centre staff give guidance in scoring when an assessment is reviewed. A more intensive academic staff development programme is being planned for the next academic year.

#### Pre-tests

Since CBA is a new experience for many students also, particularly in their first year of study, a number of tutors have asked for pre-tests to be set up so that students may familiarise themselves with the style of questions that they are likely to encounter. The Centre has completed a number of these tailored to specific disciplines, however the intention is to produce a generic pre-test covering all TRIADS question styles for student familiarisation purposes in order to rationalise the duplication in this area.

#### **Delivery issues**

TRIADS assessments can be delivered from executables over local area networks or through a web-browser for intranet and internet applications via the Authorware Web Player plugin.

Where formal summative assessments are to be delivered via the local area network they are completed under standard examination conditions with academic invigilators and at least one member of staff who is familiar with the technical aspects of assessment delivery. The assessment is started via a special network account that will only run the assessment software and prevents access to other packages or the location of results files.

Assessments delivered through the web browser have been predominantly used for open-book assessments that have replaced traditional coursework assignments. Formative assessments have similarly been delivered via the web browser with the system maintaining a record of student performance so that tutors can evaluate any problems as they arise.

For all client machines within the university network, results can be returned to any network server using UNC path conventions. For assessments taken outside the university, results are returned via FTP. Additionally, TRIADS will send results to self-refreshing web pages so tutors can view the progress of their assessments in real time from their own machines.

Delivery issues of concern encountered through the year have mainly centred upon speed and reliability of the network coupled with promotion of efficient communications between the Centre and key staff of the university Computing Services Department.

#### Efficient communication with support staff

Rapid access to key staff in Computing Services has been paramount in ensuring the reliability of delivery. The dates and locations of all computerbased assessments are published on a continually updated web page accessed by key staff in computing services and technicians responsible for the rooms used for delivery so that assessments are not disrupted by server or machine maintenance etc. All machines are thoroughly checked prior to delivery of formal assessments and sufficient excess machines are reserved so that assesses may be rapidly transferred in the event of an irretrievable crash.

#### Speed and reliability of networks

Substantial problems were encountered with the login speed and reliability of the network during the period of heavy usage at the end of the first semester. The fact that summative assessments were being run across the network acted as a spur to resolve the causal problems in the network perhaps, more rapidly than would otherwise have been the case.

These problems have not recurred at the end of the second semester even though network traffic has been similarly high. However the concentration of coursework assessments towards the end of the semester results in very intense use of the network at a time when its reliability for computer-based assessment is most important. The slow network response at such times has necessitated an increase in the FTP timeout from the twenty seconds which is adequate at normal times to over two minutes at busy times in order to prevent the appearance of disturbing 'data not saved, please retry' messages during assessments. The problem of clash between the timing of computer-based assessments and high network traffic could be resolved if tutors could be persuaded to run their assessments during the examination period itself when its use for the completion of computer intensive coursework assessments is generally complete.

## Multimedia performance

Some assessments this year have used video clips as an integral part of the exercise. These have been generally successful when run from a local area network server in terms of speed and reliability but performance can become erratic when delivered through a web browser at times of high network traffic. Delivery of substantial amounts of video material via modems is not really practical in terms of download times and supply of the multimedia resources for an assessment by CD would be preferable.

QuickTime movie files gave the smoothest performance during web delivery but are not really suitable for frame by frame or reverse play access in the Authorware<sup>TM</sup> environment at present. Where fine control is required then we have resorted to bitmap sequences as the best solution but with the disadvantage of additional file size and slow initial download to the browser. Where performance is critical, as for example in a timed summative assessment, it is apparent that delivery from an executable on a local area network server gives the most reliable solution.

## Installation of browser plugins

All student machines maintained by the Computing Services Department at the University of Derby have the Authorware Web Player, Quicktime player, Flash and Shockwave plugins installed as standard. However where students are running assessments from home they may have to install these themselves. Instructions and links to download sites are always included in the start-up page for each assessment but there have been a number of cases where students have experienced difficulty. The amount of internet time required for download via a modem can be large and a more efficient way of delivering the plugins will be investigated for the next academic year.

Part-time students attempting to access assessments from their place of work may also have difficulty in running them because of company restrictions on the installation of plugins for security reasons.

# Results filing and processing issues

TRIADS allows tutors to define up to three types of results files for any assessment. Normally these will take the form of:

- one summary file containing the names, sign-on times, time taken and total score for all assessees;
- one file containing details of individual question scores for all assessees;
- an individual results file for each assessee that contains full information on the assessee's performance on each question

including the score, question asked, the correct answer and the answer given by the assessee;

• an index file containing the filenames of the individual results files.

A typical results filing set-up for assessments delivered within the university network has been that all the files listed above would be returned to:

- a local area network server using UNC path conventions;
- the university web server or another web server as html pages via FTP;
- an additional ftp server as ASCII files via FTP.

This backup arrangement has meant that no results-files have been lost this year during filing operations. Next year it is intended to set up a dedicated assessment delivery server system with built-in backup.

For those assessments delivered outside the university network, results are returned to the web server and another FTP server via FTP. TRIADS has the capability to send results to an additional FTP server if required and an e.mail function will be added to give additional backup next year.

The slow network response at such times has been a problem towards the end of each semester as outlined earlier. This has been overcome by increasing the time-out for filing operations but filing error warnings or lengthy filing operations can be unnerving for students who are anxious that their results are saved.

The facility to configure an assessment to produce self-refreshing web pages of results has been popular with tutors and it is very useful for staff of the Centre to be able to monitor the progress of a test and filing operations.

Staff of the Centre review the results with tutors and highlight those questions that have facility or discrimination indices that are abnormal. At the end of the academic year, the scoring derived from each question will be reviewed and the question performance entered onto its Question Definition Form to allow a more informed selection of questions for future assessments.

The importance of detailed review of the individual result files for dyslexic students has been highlighted by some classical examples this year. Final scores were raised by up to 11% for an assessment containing a number of text-entry style questions. In one example, a question with seventeen text entries was scored at 12% by the system but revised to 94% on detailed inspection. In this case, the majority of answers were phonetically correct but there was no way in which the tutor could have predicted their spelling even by the use of wildcard characters.

The TRIADS system will optionally save all the answers submitted for a particular question in a single file for that question. This facility is essential for text entry questions that are being used for the first time when it is often difficult to predict the full range of scored responses. Continued use of this

facility allows the question to be modified after each run to enhance the reliability of automatic scoring.

## Types of assessment and distribution of marks

In previous years most computer-based assessments at the University of Derby were delivered under formal examination conditions, even when they were replacements for traditional coursework assignments. This year some of the more established assessments have moved to open-book, time-limited tests taken within a specified period. The comparative distributions for two such modules are displayed in Figure 1.



Figure 1 Comparison of the marks distributions for two modules that have changed from formal summative assessment (1999) to open-book time limited assessment (2000)

Overall the open-book assessments exhibit a higher standard deviation and appear to enhance the grades for students as might be expected. The questions delivered in both types of assessment were similar although not identical. It is a matter of academic judgement as to whether this change in score distribution is acceptable or whether different styles of questions may be required for examination and open-book type tests.

A more detailed investigation of results distributions is planned when all results for this year are available.

## Feedback

All tutors are encouraged to complete a Tutors Review of the performance of their assessment to enable refinement of procedures and questions for subsequent use. Any defects in question performance or ambiguities reported by students are adjusted in the master coding for the question concerned immediately the assessment has been completed.

Questions are stored in Authorware<sup>™</sup> archive files for each area of assessment. The archive files contain a question player that will display any of the questions for editing or review purposes.

## **Student Perception**

The response of students to this form of assessment is generally favourable although there are inevitably those who prefer traditional modes of assessment. However these are in the minority and the responses vary according to the difficulty of the assessment as would be expected. Assessment evaluation forms are built into the TRIADS engine and at the time of writing the results are still being collected and will be reported elsewhere.

There is a general mistrust amongst students of assessments that employ any form of negative marking. It is unusual for TRIADS assessments to be delivered with negative marking that is carried over from question to question unless the test contains a substantial number of multiple-choice questions with limited numbers of distracters. However negative scoring within questions, particularly open ended multiple-response and move object style questions, is commonly used in order to reduce the guess factor. Any negative scores resulting from such questions are zeroed at the end of the question and not carried forward to the final score.

In order to facilitate an understanding of, and confidence in, the scoring mechanisms it is our intention to develop a set of web pages that explain the scoring methods used. These pages should also be useful for tutors when designing new questions.

## Range of question styles used

Assessment systems built on authoring software are less restrictive of question design than proprietary systems. TRIADS can already deliver over thirty different question styles but modification of these is required from time to time. Tutors new to this form of assessment have tended to concentrate on the more traditional styles of multiple choice, multiple response and move object styles. However some existing question styles have been developed further.

Space precludes the description of all of these but two examples are shown in figures 1 and 2.



Figure 2 Multiple draw arrow question activated by polygonal hotspots

Figure 2 shows a multiple draw arrow style of question. The user is required to draw an arrow in each of a number of irregularly shaped areas. Arrows may be drawn as many times as required by the user within each of the areas until they are satisfied with their position. The question calculates the azimuth of the arrow assuming that the top of the screen is zero (north on maps). The tutor may define an azimuth range, within which a maximum score is attained and a score limit range within which scores are allocated on a sliding scale according to how close the submitted azimuth is to the maximum scoring range. This question template may be used to draw straight lines or forward, backward or double-ended arrows within up to twenty-six irregularly shaped areas. Use of this question in double-ended arrow mode could be ideal for a matching pairs type of question.



Figure 3 Inclined sliders

Figure 3 illustrates a question from a first year Business Studies assessment (courtesy Derek Fry) where the user is required to move two sliders across a graph to the correct positions. As with the draw arrow question, the scoring is graduated so that the nearer the user positions the line to the maximum scoring position the higher the score.

# Developments

By the beginning of the next academic year, the TRIADS system will have interactive wizards for system configuration and for a number of question styles that should speed up assessment production and make the upload and configuration on network servers substantially easier. It is to be hoped that the enhanced ease of question generation provided by the wizards will encourage more tutors to input some of the simpler question styles for themselves.

Some question styles have been re-coded to facilitate the generation of questions with combinations of interaction types (e.g. polygonal hotspot + text entry) and enhanced feedback options for formative use. Some of the standard question styles will have their coding bound into the TRIADS engine in order to reduce file size.

Data on question performance in terms of facility and discrimination will be built into each question delivered this year and the TRIADS engine will be modified to allow tutors to see the expected average score and level of discrimination for an assessment at the compilation stage.

## Conclusions

This first academic year of operation of the new Centre has inevitably encountered some problems in the delivery of assessment on a universitywide basis. Most of these were associated with network functionality and speed during the first semester and almost all have been overcome. A good working relationship and constant dialog between the Centre and the Computing Services Department of the university has been established.

Scheduling an even work-flow for assessment production is a major goal for the next academic year coupled with a wider dissemination into more disciplines. However the provision of a central department for interactive assessment development has already promoted the implementation of computer-based assessment across a wider range of disciplines at a faster rate than would otherwise be the case.

Academic staff development in the area of assessment and question design will be high on the priority list. Particularly with respect to balancing the content of questions to match the learning outcomes and to promote an understanding of the scoring and weighting within questions.

TRIADS system developments planned for the coming year will enhance the production rate and encourage tutors to become more directly involved in question production.