

## Summary

This is a summary of the research report for the first phase of the PASS-IT Project. It seeks to extend the growing evidence of the role of the computer in both formative and summative testing [1 - 5]. Working with the Scottish Qualifications Authority over the last three years, researchers at Heriot-Watt University produced some preliminary studies from approximately 100 candidates at four centres. These results were obtained in Higher Mathematics tests and sought to evaluate the role of the medium in traditional tests [6]. In addition, the studies investigated how partial credit might be incorporated into automatic summative examinations [7].

In this present study, traditional tests in Chemistry, Computing and Mathematics were chosen in an extensive study over ten centres with more than four hundred candidates. The research questions to be addressed vary from subject to subject and are summarised below.

Moving from a paper examination to one on computer involves several changes. For example, the questions may have to be reworded and/or rephrased and the answers have to be entered via a keyboard or mouse instead of by pen on paper. So, in Chemistry and Computing, in line with the earlier studies in Mathematics [3,6], the following questions were posed:

- Does changing the medium for test delivery have an effect on the marks obtained?
- Does changing the wording for ICT delivery have an effect on the marks obtained?

Using a general linear statistical model which takes account of ability, gender, school and class, no evidence of a medium or re-wording effect was found in Chemistry and Computing. In Chemistry, there was a small gender bias with slightly higher average marks for male students.

In Higher Mathematics, the earlier investigation [7] required further study. For instance, in a paper examination partial credit can be gained for answers which are incorrect but contain some of the correct ideas. This is normally awarded on the basis of the number of learning outcomes that have been satisfied. So, the following questions were asked:

- How does the student performance on paper where partial credit is awarded compare with performance on an ICT exam in which optional steps are available (but where a penalty of marks lost for choosing steps may accrue)?
- How does the student performance on paper where partial credit is awarded compare with performance on an ICT exam in which the student is given immediate feedback in the form of ticks and crosses to indicate if right or wrong and, if wrong, given the opportunity to resubmit the answer as often as desired?

There was no evidence of a difference between these forms of the Higher Mathematics National Assessment Bank (NAB) unit 2 test where again ability, gender, school and class were variables in the statistical model. There was weak evidence of differences on a couple of individual questions due, in part, to an inaccurate distribution of marks in the automatic marking scheme and some follow-through problems. Not surprisingly, there was some evidence that students in the middle range of ability took advantage of steps in a few of the questions.

In Advanced Higher Mathematics, the same issue of partial credit occurs as in the Higher Mathematics but the questions are at a more advanced level giving rise to more complex issues. The research question in this part of the study was formulated as follows:

- How does the student performance on paper where partial credit is awarded compare with performance on an ICT exam in which steps are available to the student (but at a penalty of losing marks)?

The Advanced Higher Mathematics NAB unit 1 was split into two half tests and students in two schools took one of the half tests on paper and the other half test on computer. The statistical tests revealed weak evidence of a difference between the two forms of the test but again only in one question where follow-through errors occurred for those taking the computer test. There was weak evidence of a gender effect with females scoring more marks on average.

The research has sought to identify issues for questions where appropriate, for both those that cause no difficulty in translation from paper to computer and those that do. There has been an emphasis on special educational needs throughout the study and experiences have been reported as appropriate. Efforts have been made to produce tests on computer in which the user can change font in size, style and colour with, in addition, a choice of background colour. Also, the tests can be used with the screen reading package JAWS to help blind and partially sighted users.

Finally, PASS-IT has been able to provide a broad study ranging from answers in the form of mathematical expressions to those involving a free text response. In the former case the technology developed at Heriot-Watt University has been employed (see [8] for more details) and in the latter case the services of the software firm Intelligent Assessment Technologies has been used [9], based on their own technology. The main aim of this part of the work was to investigate present and future capabilities of automatic marking of free text responses. Intelligent Assessment in a separate report [10] addressed the ability of their technology to mark electronically a range of short free text responses in Mathematics, Chemistry and Computing. Their report predicts the number of student responses needed to provide more accurate automatic marking of free text responses across this range of subjects. Their engine experienced some difficulties with mathematical notation and the more open-ended questions in Computing but coped very well with free text responses in Higher Chemistry.

All the data from the project is being analysed before conclusions are drawn. But, phase 2 will attempt to extend the range and variety of phase 1 through work in Intermediate English, Higher French, Higher Music as well as Access and Intermediate Mathematics. Again, special educational needs will be built into any trials at the outset. In addition, responding to enthusiasm from teachers, PASS-IT is embarking on an alternative evidence trial in two Edinburgh schools in Higher Mathematics NAB unit 2.

It is appropriate at the end of this short review of phase 1 of PASS-IT to thank all the teachers, lecturers and technical staff in schools and colleges who welcomed the PASS-IT team and helped the collection of data in the experiments described in the full report. It is also a pleasure to report on the excellent team work, good humour, common sense and hard work from all those involved in the PASS-IT Project from Advisory Group, Project Board and its many team members.

## References

- [1] Beevers, C, Foster M and McGuire, G, (1989) Integrating Formative Evaluation into a Learner Centred Revision Course, British J of Educ Tech 20, 115 - 119.
- [2] Harding, R and Quinney, D, (1996) Mathwise and the UK Mathematics Courseware Consortium, Active Learning 4, pp 53 - 57.
- [3] Beevers, C, McGuire, G, Stirling, G and Wild, D (1995) Mathematical Ability Assessed by Computer, J Comp Educ 25, 123 - 132.
- [4] Greenwood, L., Cole, U.M., McBride, F.V., Morrison, H., Cowan, P.& Lee, M. (1998) Can the same results be obtained using computer-mediated tests as for paper-based tests for National Curriculum assessment? Int Conf Mathematics / Science Educ and Tech, Vol. 2000, Issue. 1, 2000, pp 179-184.
- [5] Beevers, C, Youngson, M, McGuire, G, Wild, D and Fiddes, D (1999) Issues of Partial Credit in Mathematical Assessment by Computer, Alt-J, 7, p26 - 32
- [6] Fiddes, D, Korabinski, A, McGuire, G, Youngson, M and McMillan, D (2002) Are Examination Results affected by the mode of delivery? Alt-J, 10 p60 – 69.
- [7] McGuire, G, Youngson, M, Korabinski A and McMillan, D, (2002) Partial Credit in Mathematics Exams: A Comparison between Traditional and CAA, Proc Sixth International CAA Conference, Loughborough University pp 223 - 230.
- [8] <http://www.calm.hw.ac.uk/cue.html>
- [9] Mitchell, T, Russell, T, Broomhead P and Aldridge, N, (2002) Towards the robust computerised marking of free text responses, Proc Sixth International CAA Conference, Loughborough pp 233 - 249.
- [10] Mitchell, T, and Aldridge, N, (2003) Computerised Marking of Free-Text Responses : PASS-IT Project Report for the SQA.