A PILOT SYUDY IN WHICH GEOGRAPHY STUDENTS ASSESS THEIR LEARNING USING COMPUTER-MARKED TESTS

James Hogg

WORKING PAPER 95/9

SCHOOL OF GEOGRAPHY • UNIVERSITY OF LEEDS

A pilot study in which geography students assess their learning using computer-marked tests

James Hogg

School of Geography

University of Leeds

LEEDS LS2 9JT UK

April 1995

1. Introduction

A new concern of lecturers in higher education in UK is how to provide effective and continuous guidance to first year undergraduate students as class sizes increase and resources decrease. With almost 400 students in first year modules in geography at the University of Leeds, lecturers have become acutely aware that traditional methods of teaching, assessment and giving guidance to individual students are no longer practical. The time and effort involved in providing constructive criticism and personal tuition to each student on a regular traditional basis is becoming an intolerable burden on staff resources. The great temptation is to reduce assessed work and personal tuition in order to preserve time for research. If this were to happen however it would have unfortunate consequences for students learning and would undoubtedly dampen their interest, enthusiasm and motivation to excel in geography.

Most students welcome comments from lecturers about their learning activities and performance, partly for diagnostic purposes and partly as a stimulus to work (Lowman, 1987). They want constructive criticism of their work, frequent guidance on how well they are doing and reassurance about their progress. Lecturers recognise and accept this. For that reason, they will continue to provide such criticism, guidance and reassurance. The debate is not about whether or not they should do this but about how they should do it. The problem is how to provide objective, efficient and continuous guidance to students as class sizes increase and resources decrease. As Entwistle (1992) stresses in his review of the relationship between teaching and

learning outcomes in British higher education, there is substantial evidence that teaching can be adapted to facilitate desired learning outcomes but academic staff must be prepared to adopt newer methods of teaching, have a positive attitude towards them and be well-versed in the principles underlying them.

A novel method of solving this problem is being pioneered by colleagues in Biochemistry and Molecular Biology at Leeds (Booth, 1994). With considerable knowledge and experience of using multiple choice questions (MCQs) in medicine and related fields for assessing students' learning, they have been experimenting with a method in which small-group tutorials for large classes are replaced by "a computer-based support targetting system"(Hassall, Turner & Wood, 1987; Booth and Wood, 1989). In a recent paper describing the method, Booth (1994) explains that every two weeks, the students complete a short multiple choice test. The questions relate to the topics being taught in lectures. The tests are compulsory. After students take the test, the lecturer explains each answer, thus providing immediate feedback to the class. The students' average scores comprise a component of the end of semester grade. The tests are marked using a computer and the students scores are published within a day taking the test. In the week following the test, those students who have low scores are contacted individually and offered help. Thus, there is a means of targetting and giving help those students who need it most.

The method offers scope for use in large undergraduate classes in geography. Though there are dangers in assuming that what works well in one subject will necessarily work in another and that technology-based assessment will solve all problems, the method is attractive as a complement to traditional forms of assessment in geography.

In order to assess its potential, a pilot study was undertaken in which second year students in geography sat multiple-choice tests in lectures and practical classes. This paper examines the context, results and issues surrounding the use of such tests in geography.

2. Managing assessment using multiple choice questions

Teaching and assessment are closely interrelated in higher education. Both have lately undergone critical scrutiny and reappraisal partly in response to the HEFCE Quality Assessment of Teaching exercise and partly in response to the increased number of students in geography modules. The aims, objectives and methods of teaching have as a result been reviewed and in many cases overhauled. The outcome has generally been to improve standards of teaching and learning in geography. It has clarified aims, sharpened objectives, improved methods and specified outcomes in terms of what students should gain from a module. One of the crucial criteria in judging the quality of teaching is however the method of assessing students taking a given module (Brown and Knight, 1994).

The assessment of students on a course module serves a variety of functions (Brown, Rust and Gibbs, 1994). These typically include *inter alia* generating motivation in students, stimulating appropriate learning activities, providing feedback, ranking and grading of students and providing evidence for quality assessment of teaching and course evaluation. Other functions can be added which are specific to geography. Taken altogether, the functions can be divided crudely into those that are formative in

that the student benefits directly from the learning process associated with these and summative in that the assessment is for the benefit of the lecturer and the institution. Though there is clearly much overlap between the two, formative methods improve the learning experience of students.

Both formative and summative methods of assessment become a heavier burden for lecturers as class sizes increase. Formative methods are generally much more demanding of a lecturer's time. As a result, lecturers need to explore new methods of formative assessment, preferably those which are economical in the use of resources, objective, accurate, practical to implement and efficient (Gibbs et al. 1992).

The Multiple Choice Question (MCQ) test has been widely used in North America though it is a relatively new form of examination in this country. It can be used with large classes, lends itself to computer marking and gives standardised results which are free from personal bias. Chiefly because it is perceived as being fair, it has gained acceptance at all levels of education. Hassall et al (1987) stress that the MCQ test is probably here to stay for several reasons, "many of them economic rather than academic". Used widely in medicine and related disciplines, the MCQ test has become popular as one of the methods of assessing knowledge and progress.

The MCQ test is not without its critics who rightly ask what it was assessing. Whatever it was assessing, it was not the same as traditional methods of assessment in geography. The acquisition of knowledge in geography was but one aspect of the study for a university degree.. What MCQ tests do not assess are a student's ability to

select and synthesise information, to organise their ideas, to write essays, to prepare critical reviews and to write project reports. In view of this, lecturers would be unwise to base any assessment of students solely on an MCQ-test. Students must continue to be assessed by the methods which have been used successfully in the past and which assess different academic and personal qualities. The MCQ test has nevertheless a role to play as we diversify our methods of teaching and assessment in higher education and as we attempt to cope with the increase in class sizes in 1990s. MCQ tests provide another arrow in our quiver for assessing a student's knowledge of information, for monitoring progress and, to a more limited extent, for gauging understanding of concepts, theories and methods

3 Computer-based marking

The MCQ-test aims to provide students with frequent short tests and feedback for self-diagnosis of their progress in learning throughout a course module. If applied systematically throughout a degree scheme, they would be built into the course, take place every fortnight, give immediate and full feedback to each student about performance and act as a stimulus for students to work continually at their studies. The importance of generating and maintaining motivation in students, by providing frequent tests and feedback, lies at the heart of this formative rather than summative approach.

MCQ-tests are designed to test specific knowledge, understanding, comprehension and analyses and a broad range of mathematical, statistical and computing skills.

They can be based on several types of question, probe different levels and types of learning and understanding and give an insight into what students have learnt during the period immediately prior to the test. Various types of teaching resources can be used during tests such as texts, reports, statistical summaries, samples, models, maps, charts, photographs and remote sensing images. Each test typically lasts for twenty minutes, takes place in a lecture theatre or practical class and becomes a routine part of a module.

For an MCQ-test, each student is provided with a question paper and a printed computer card. The printed card has space for students to write their name, class and test number. This is purely a precaution in case cards get mixed up and is not read by the computer. Each card is divided into two sections. One section has a matrix of 10 columns by 9 rows of printed boxes, each column numbered from 0 to 9, in which students identify themselves by marking their University registration number. The second section contains a matrix of five columns by 50 rows of printed boxes. The five boxes are labelled A, B, C, D and E and rows are numbered from 1 to 50. This means that tests can have up to 50 questions, each with five possible non-exclusive answers. One or more answers must be correct. Students mark their answers with a soft lead pencil by marking the appropriate boxes.

At the end of the test, after the computer marked cards are collected, lecturers not only give the correct and incorrect answers but explain them thereby providing immediate feedback to students.

A card reader attached to a personal computer is used to read data from the cards. Each card is passed through the reader twice in order to identify any cards which have not been properly marked and may cause mis-reading errors. Negative marking is used for wrong answers but the minimum score for each question is zero. Hence overall negative scores are impossible (Booth 1984). Using Microsoft Windows software specially written by Booth (1994), the cards are processed typically within hours of the test. The University registration number and name of all students registered on the course module is down loaded from University Records into a file in Lotus 123.WKS format. The lecturer fills in a Master Card with the correct answers to all questions. The Master Card is read first followed by student test cards. The program matches student numbers with names and produces results for each student and for the class. Results can be manipulated in any spreadsheet package such as Lotus-123 or Excel. An analysis of the collated results can be generated in terms of facility index (measures difficulty), discrimination index (measures confusion) and selection frequency for each question (Brown and Pendelbury,1992). As a rough guide, to read and process 100 cards takes about an hour.

4 Results

Some 70 students taking level 2 modules in "Climate and Remote Sensing" and "Physical Laboratory Practical" in Geography were given MCQ-tests. The 20 minute tests contained 20 questions. One group was given the test in a lecture theatre and the other in a practical class. Results showed a range of marks from 72 to 25%. Students with the highest marks typically got most correct answers but there were diverse ways

of doing well. For example, some questions had four correct answers but students only selected one as being correct. Therefore they missed opportunities for more marks. Some students got all the answers correct in such questions but had more wrong answers elsewhere. The possible permutations means that students had not only to get a right answer but all of the right answers to do well. In other words, they had to know about the subject.

Results from the pilot study show mutual benefits for students and staff. Students learn immediately after a test where they are ranked in their class. They like to know where they stand, and, with a series of such tests at fortnightly intervals throughout a course module, they can monitor their level of understanding and progress. If they are weak in certain areas, they can decide themselves to devote more time and effort to make good any shortcomings. Those students who study as a fear of failing learn differently than those who learn to excel. The effect of tests will be different in individual cases but results suggests that a majority strive to do better.

Staff can use results in a variety of ways. They can readily determine which topics were taught well and understood by a majority of students and which need to be revisited and explained again perhaps in a different way. Given that the lecture continues to be the main vehicle of instruction in most universities in the UK, the need to make sure that it is stimulating and effective in communicating key concepts, theories and methods in geography remains of paramount importance. Tests provide feed-back to the lecturer at regular intervals throughout a module about how perceptions of teaching are translated into student learning and understanding.

- Lecturers can see immediately which students are in need of help and can arrange compulsory remedial tutorials to try and help students who are weak to keep abreast of work and motivated.

Results from the pilot study confirm what educational experts have long been telling us. Geography students like to have specific feedback about their progress in understanding topics in lectures. They like to know where they have strength as well as where they are weak. They like to know where they stand and where they are ranked in their class. They like to monitor their progress in learning, to relate their personal learning activities to the outcomes of tests and to relate other personal matters such as missing lectures through illness or other causes to test results.

Students accept computer-marked tests as a way of assessing and monitoring their learning. They view them as being objective and fair and value the immediate and accurate feedback. Most have experience of using computer-marked tests from school. They recognise that computer-marked tests are an efficient way of assessing their learning. They felt that MCQ-tests were a good way of assessing their learning. Some complained however that the MCQ-tests were only assessing knowlegde, that they needed more time to prepare themselves for such tests and that the questions were far too specific and difficult. Others felt that the questions were too easy!

The setting of an MCQ-test is a skill which requires clear objectives, ingenuity and time (Hassall, Turner & Wood, 1987). There are several generic types of multiple-choice questions for eliciting different types of knowledge, understanding, analysis,

mathematical or computational skills. The setting of MCQs should be a team effort. It is a skill which demands a knowledge of a subject and of MCQs, as recent case studies have shown (Bull, 1993). The investment of time and effort in setting MCQs is significant but can be recouped most effectively if the MCQ papers, distributed before each test, are collected in immediately after the test is finished. In this way, the same MCQs can be used over a period of years.

5. Conclusion

MCQ-tests have much to commend them as a way of allowing students to monitor their own learning and progress on modular courses in geography. They can be given in ordinary lecture rooms, to very large classes and can provide objective, accurate and fair assessments as well as providing immediate feedback to students. The resources required to introduce such tests are modest: a card reader attachment for a personal computer, computer software and printed computer cards. As far as students are concerned, all that they require is a soft lead pencil. The computer marked cards provide a permanent record of student progress on a course. They can of course be used for other purposes such as course evaluation on a year by year basis, assessment of lecturers and teaching quality audits (Houghton and Booth 1994). Though many lecturers have misgivings about assessing students of geography using MCQs, the argument should be focused on giving continual feedback to individual students as a means of self-diagnosis of their learning, identifying strengths and shortcomings and motivating students to work in the areas where they need to in order to cover all the material in a course module. The argument should also be focused on introducing

greater diversity into the assessment process with a method which has specific advantages for large classes. If introduced systematically throughout a department, the method has major benefits both for staff and students.

There has been much discussion in the literature about computer-based learning (CBL) and automated assessment of students while they work at computers. Though technically feasible and worthwhile, the resources are simply not available, and appear unlikely to be available for some time to come, to contemplate such an approach for the large classes in level 1 geography modules. Computer marked tests offer an alternative which is economical, labour saving and highly cost-effective as a formative method of allowing students to monitor their learning. Whatever doubts geography lecturers may have about the use of MCQ-tests as a means of assessing students learning and progress, and there are substantive and valid arguments against the indiscriminate use of such tests, computer marked tests offer a means of self-defence for lecturers who are faced with the significant changes and pressures affecting higher education in UK in 1990s.

6. Acknowledgements

I would like to thank Dr Andrew Booth, Biochemistry and Molecular Biology, for his advice and help in processing computer-marked cards. I would also like to thank Dr Frances Drake and Dr Richard Smith for help in conducting the pilot study.

6. References

Booth, A.G. (1984) The use of microcomputer equipment with a marked sensed card reader for student assessment. In Wood, E.J. (Ed) *Microcomputer in Biochemical Education*, 181-193. Taylor and Francis, London.

Booth, A.G. (1994) Targetting student support in large first year classes. *Biochemical Education* Vol 22 199-200

Booth, A. G. & Wood, E. J. (1989) The microcomputer in biochemical education. *Progress in Medicinal Chemistry*. Vol 26 323-354

Brown, S. & Knight, J. (1994) *Assessment in higher education.* Kogan Page, London.

Brown, S., Rust, C. & Gibbs, G. (1994) *Strategies for diversifying assessment in higher education.*. The Oxford Centre for Staff Development, 52 pp.

Brown G A & Pendlebury M (1992) Assessing active learning. CVCP/USDU, University of Sheffield, 62 pp.

Bull, J. (1993) *Using technology to assess student learning.* TLTP Project ALTER, UK Universities' Staff Development Unit and the Universities of Kent and Leeds, CVCP/ USDU, University of Sheffield, 62 pp.

Gibbs, G. et al. (1992) *Assessing more students*. No 4 in Teaching More Students series. Oxford Centre for Staff Development, Oxford Brookes University, 45 pp.

Entwistle N (1992) The impact of teaching on learning outcomes in higher education. CVCP/USDU, University of Sheffield, 59 pp.

Hassall,H. Turner, A.J. & Wood, E.J. (1987) *Multiple Choice Questions in Biochemistry*. Churchhill Livingstone, Edinburgh.

Houghton JD and Booth A G 1994 Making your mark with course questionnaires. *Biochemical Education* 22(4): 210-203

Lowman, J. (1987) *Giving students feedback* In M G Weimer (ed) Teaching large classes well. Jossey-Bass, London.