Using Technology to Assess Student Learning

The UK Universities' Staff Development Unit and the Universities of Kent and Leeds

TLTP Project ALTER

December, 1993
ISBN 1 85889 091 8

Universities' Staff Development Unit,
Level Six,
University House,
Sheffield S10 2TN.
Tel: 0114 225 5555
Fax: 0114 225 2042

The UK Universities' Staff Development Unit
and The Universities of Kent and Leeds
Using Technology to Assess
Student Learning

Joanna Bull
Research and Administrative Officer
Project ALTER

Contents

Preface

This paper outlines some of the ways in which technology can be used in the assessment of student learning. The two main uses of technology in the assessment process are identified as: using technology to perform the actions and cognitive processes a tutor would employ to arrive at a grade, and using technology to aid the assessment process. The paper concentrates on the use of technology to grade students' work, with the emphasis on the use of technology to enhance the efficiency of assessment procedures.

There is currently much development in harnessing technology to teaching and learning. It would be near-sighted therefore not to consider its use in the assessment of students, who may spend much of their university lives using technology to enhance their learning.

Approaches to assessment, from essays and MCQs to self and peer assessment are covered. Guidelines are given as to how technology might be used in conjunction with these methods. Case studies explore the practical implications of technology-based assessment systems, and provide readers with a point of reference for further investigation.

Chapter Six explores some of the uses of technology as a tool in the assessment process. The technologies discussed include video, computers and multimedia. Case studies highlight the advantages and disadvantages of these methods over traditional assessment methods.

The penultimate chapter investigates the uses of technology to manage assessment information. It emphasises institution-wide efficiency gains as a means of managing rising student numbers and modularisation.

The final chapter offers some guidelines for those who intend to use to technology to assess students' learning. It highlights some of the pitfalls and provides words of caution for those intending to embark on technology-based assessment. It encourages careful consideration of a variety of factors and highlights the advantages of technology-based assessment. The chapter ends with speculation about the future. It is highly recommended that lecturers read this booklet in conjunction with Assessing Active Learning, Brown and Pendelbury (1992), available from USDU. This gives information on a wide range of assessment techniques in higher education. I am greatly indebted to the academic and administrative staff at a considerable number of universities who kindly contributed to this publication.

Chapter One

Introduction

Traditionally assessment in higher education has been about examinations. Recent changes in the nature of higher education have led to changes in the ways students are assessed. As student numbers increase staff must inevitably spend more time on student assessment. No longer can many staff claim to know all their students well. Their time is consumed with marking. Modularisation, audit and quality assessment may necessitate further changes, which may in turn affect the nature of assessment. It is understandable that there is concern about the quality of assessment. There is a critical point, an amount of marking beyond which everything deteriorates - feedback, accuracy, speed, judgement and health. It is no wonder that alternatives are being looked for.

A major influence in recent years has been the Enterprise in Higher Education initiative (EHE). Many of the EHE programmes concentrated their efforts on the development of transferable skills, curriculum design for active learning and innovative teaching methods (Entwistle, 1992). A range of skills are being learned in higher education and there is a need for a new range of assessment tools and methods to be developed, to respond to this shift in emphasis.

Many of the EHE programmes have contributed to the development of technology-based course materials. Computer-aided logic teaching in the Humanities at the University of Ulster, and computer-based learning to demonstrate psychological bias in decision-making in Business and Economic Studies at the University of Leeds are two examples only. Many subject areas now use one or more of the following: multimedia applications; hypermedia; computer-assisted learning materials; video and so on. Technology-based innovation in teaching methods and the presentation of course materials has in some cases led to technologically-based assessment methods and tools.

The purpose of this paper is to provide a snapshot of ways in which technology can be and is being used in the assessment of students. It will not present a complete picture. The pace of technology development and innovation in higher education is such that change takes places rapidly, and some schemes and projects are disregarded in favour of more up-to-date technology before they have really taken hold.

Introduction

This paper will be of interest to those who want to find out more about the uses of technology in assessment. It has three purposes:

- to disseminate ideas and schemes which may prove useful to those wishing to investigate the use of technology to aid student assessment;
- to provide examples of good practice and innovation in technology-based assessment methods and procedures;
- to identify the pitfalls of the 'technology trap'.

It is hoped that this publication will act as a reference point for those wishing to learn about, or experiment with, the use of technology to assess their students.

The examples used in this publication were collected from a range of disciplines and higher education institutions. For further information about a particular example, references and a contact list are given at the end of the paper, to which the number in brackets after each example/reference corresponds.

Chapter Two

Definitions

Before getting involved in the detail of the ways in which technology can be used to help assess students, it would be helpful to provide working definitions of the terms 'assessment' and 'technology'.

Assessment

Assessing a student involves taking a sample of what he or she can do, and drawing inferences about a student's capability from this sample. It should be an appropriate and relevant sample which represents what should have been learnt. However it will be only a sample - it is not possible or even necessary to assess every skill, concept and ability that has been taught.

Summative assessment involves methods designed to establish what a student has achieved at the end of a unit or course. A final mark or grade is awarded.

Formative assessment involves methods designed to establish what progress a student is making during learning and to enable giving feedback on it.

Technology

For the purpose of this paper 'technology' is defined as methods and materials which are relatively machine-dependent. It includes the following: computer software packages; computer-assisted learning (CAL); computer-based learning materials (CBL); networks; video; hypertext; hypermedia; simulation; multimedia; scientific visualisation and virtual reality.

Although the present state of student assessment through technology does not include all these applications, the most popular technology is computer software. However there is increasing use of multimedia and hypermedia applications in teaching and assessment, and advancement in the fields of virtual reality and scientific visualisation which may result in the use of such applications in the future.

Readers may be familiar with CAL and CBL materials, and may even have used or created some. The fields of hypertext, hypermedia, simulation and multimedia may not be so familiar, and so I outline brief definitions for each:

Hypertext: The ability to make multiple layers of text appear on the screen through the creation of links between the texts.

Hypermedia: This technique extends the concept of hypertext to include other media such as sound and video.

Simulations: the emulation of real-life activities and occurrences using a computer with graphical capabilities.

Multimedia: (sometimes known as interactive media) is the ability to combine several different media, such as graphics, sound, video, text to create an interactive presentation. Although this is similar to hypermedia, it differs in that hypermedia is used to present large amounts of interrelated information. Multimedia is mostly used to give an animated presentation.

Assessment and technology

The uses of technology in assessment of students can be separated into:

- technology as a tool of assessment;
- technology to assess learning.

Technology as a tool of assessment is found fairly commonly in UK universities. By this, is meant the use of a technological facility to aid the process of assessment. Asking students, for example, to make a short video film to illustrate the flow of people through various routes in a National Park, and then viewing and grading it, would be using video as a tool of assessment. This could have been undertaken by setting students an essay question or asking them to make an oral presentation. Technology is used as a tool of assessment for a variety of reasons - it may prove more efficient to watch 15 ten-minute videos than 15 ten-minute oral presentations. It may be that students learn a new range of skills and gain valuable experience from making the video, or that certain skills can be assessed more effectively.

Whatever the reason, it is becoming increasingly common to use a range of technologies to supplement or replace traditional paper and pencil tests. Using technology to assess involves the use of technology to assign marks to an essay, practical or project. That is to say that, the technology performs the task, or part of the task, which would otherwise have to be done. For example, using an optical scanner to mark

Chapter Three

Multiple-choice Questions

Introduction

Multiple-choice questions (MCQs) are often seen as an efficient option when looking for a method which assesses large numbers of students quickly. They take the form of a stimulus followed by some alternative responses from which the student must pick the correct alternative or alternatives. The answers can be in the form of true-false, yes-no, or more alternatives from which students must choose or match a series of statements/reasons.

Although MCQs can provide a rapid means of assessing large groups of students, as marking time is reduced through the nature of the test, considerable time and effort is needed to prepare and design the questions. The time-saving occurs in the easy administration and marking of the tests. The re-use of items from a secure 'bank' enables further time to be saved.

MCQs can be used for both formative and summative assessment. The results can be passed to students, highlighting areas on which they need to concentrate with relevant references. Lecturers can use the results to guide them in refining lecture topics and course planning. MCQs are often criticised for encouraging guessing, though it may be considered that guessing is a useful skill and should be encouraged and measured.

The scoring system to be used should be given to the students, as this will affect their response to the test. Incorrect answers and unanswered questions may either score a minus or nought. This will affect how much guessing the students do.

The number of alternatives to a question is also of importance. As the number of alternatives increases, the effectiveness of guessing in improving scores decreases. This should not be taken as the sole reason for deciding the number of alternatives to use. Other factors come into play, such as the depth of thinking and reasoning that is required to consider each alternative, how long the students have to complete the test, and how many plausible alternatives can be created. multiple-choice questions involves relinquishing the marking of multiple-choice answer papers and letting a machine do it.

It is sometimes difficult to separate the two, but this paper concentrates mainly on the use of technology to assess. In some methods of assessment however, it may be appropriate to use technology only as a tool of assessment.

Much has been written to give guidance on the use of MCQs. Brown and Pendlelbury (1992) Assessing Active Learning is strongly recommended and gives useful detail specifically on assessment in higher education.

Multiple-choice questions and technology

The most common use of technology for assessing students is in the application of MCQs. The nature of them means that, given appropriate software, it is fairly easy to convert a MCQ test on paper to a computer without alteration to format or content.

Several options are available to creators of computerised MCQs, which are not available to those using paper and pencil MCQs.
These are:

- to allow more than one attempt at each question;
- to supply hints to the user;
- to give the student the choice to discover the answer (a ‘cheat’ key);
- to feed distracters to the student as they progress;
- to be purely for self-assessment, or also for summative assessment;
- to check whether the student has used the test, even if only for self-assessment;
- to allow the student more than one attempt even if the test is for formative assessment.

The decision to employ any of these options depends on the purpose of the assessment. Many early computerised MCQs were used for self-assessment only. It is now possible however, to guard against cheating with time checks for each answer, examination conditions, and randomly generated tests where each candidate is given a slightly different test. This allows far more flexibility in the use of computerised MCQs. This is advantageous because flexibility allows MCQs to be used for a variety of purposes - self-assessment, summative testing, and formative, informal assessment. However, this also means that the system must be carefully maintained, questions updated regularly, and for formal, summative assessment, students must be organised into formal settings and invigilated.

There are a number of ways in which technology can be used to administer MCQs. These are as follows:

- A bank of questions can be stored from which question papers are generated, and printed out. Papers can be generated as a standard test, where everyone gets the same questions, or alternatively randomly, where everyone has a different test, and tests of varying levels of difficulty can be distributed. These are given to students as paper and pencil tests.
- An MCQ test can be presented on screen which, when completed, is assessed by software. Results can be transferred to record files, printed out, and displayed to the student.
- An optical scanner or optical mark reader can be used to mark papers, usually on specially-designed forms.
- Using appropriate software results from optically-scanned papers can be collated, statistically analysed and stored.

Commercial packages

There are a number of packages available commercially which enable the generation, assessment and scanning of multiple-choice tests. A number of companies, (such as *NSCi* (1) and *DRs* (2)), produce software and data-processing systems. Some - such as *NSCi's ExamSystem* software - have the ability to link into an optical mark reader or scanner.

With such software it is possible to make a selection of items from a bank of questions chosen using certain criteria, such as levels of difficulty, type of question. Question papers are generated, the test is taken and with the use of an optical scanner, results can be achieved and feedback to the students can take place very rapidly.

Many lecturers use packages such as *Question Mark* (3) to design and administer their MCQs. *Question Mark* can create three types of question file, varying in the amount of feedback presented to the student (from none to immediate). The questions can be of nine basic types. As well as MCQs there are: number entry; word entry; gap fill and free-format, where the student's entry is compared to a correct solution through a key word search. Packages such as these can be used to present a more sophisticated assessment, but need precision in their construction.

MCQs can also be constructed using hypertext. A set of self-assessment questions has been developed at the University of Kent at Canterbury for use in the Kent Law School (4). The questions were written in hypertext, with support from the Hypertext Support Unit. Students are asked questions and given three possible answers. Using a mouse they click on the one they think is right, which then expands to reveal details about their choice. If they are wrong, they can close it up and try again.

More recently, educational institutions are co-ordinating their expertise, in the form of CBL Support Units, and Institutes of CBL, to enable a variety of academics to benefit from others' experience.

A recently developed package is *Examine: Multiple-choice and Short Answer Quiz Authoring System* (5), developed by the Bio-Informatics Research Group at the University of Nottingham, under the Information Technology Training Initiative (ITTI). *Examine* can be used to write and administer MCQs and short answer tests in a graphical environment (currently only *Windows*, although *Macintosh* is under development). It can be used for self-assessment and revision, although it is not intended to be secure, as data is stored in ASCII files for ease of accessibility. The system allows the use of random test construction from a potentially large database, flexible marking schemes, and the storage of results in a file. The system is in prototype form for discussion purposes, but a new version is currently under development. This will be a completely re-written version available on both the *Macintosh* and *Windows* platforms. (This means that it will be possible to author on one platform and deliver on the other.) The data structure will be more flexible, and free or guided browsing through a structured question-base will be possible. There will be a larger variety of question types; the question formats will be more flexible; and there may be multimedia support. It is hoped that both *Macintosh* and PC versions will be available in early to mid 1994.

Another similar package is *Question Handler* (6), developed at the Institute of Computer Based Learning, Queen's University, Belfast. It has the capacity to support Yes/No questions and MCQs, and may be developed further to include 'match' type questions.

Who uses MCQs?

MCQs are used in a variety of different disciplines, although they are probably more widely used in the Sciences than in the Arts. The factual content of subjects such as Medicine, Biochemistry and Pharmacology mean that the construction of MCQs to test knowledge and its application is somewhat easier than it would be in, for example, English Literature which may require more sophisticated MCQs. Nevertheless, MCQs have been used in the following disciplines: Information Science; Accountancy; Medicine; Management; Psychology; Orthopaedic Surgery; Mathematics; Architecture; Russian; Pharmacology; Biochemistry; Philosophy; Economics; History and Community Health to name but a few. In the majority of these disciplines MCQs can now be presented in computerised form.

Case Study 1 - Department of Biochemistry and Molecular Biology, University of Leeds (7)

The Department of Biochemistry and Molecular Biology teaches a first-year science class of about 380 students each year. The year's teaching takes the form of two semester-length courses. A systematic approach to the use of MCQs has been adopted in order to assess students' progress, identify any topics found difficult to understand and then use this information to target tutorial support rapidly.

Every week all first year students complete a short MCQ test on material presented in the lectures they have recently attended. Each question has five alternatives and students mark their answers on specially designed cards. The students undertake the test in the first twenty minutes of a three-hour practical. The cards are collected and optically scanned. It takes about one hour to scan one hundred answer cards. The mark sense card reader is small compared with the A4 readers often used, is relatively cheap (about a third of the cost of an A4 reader) and is simple to use. The cards used are printed locally at a cost of approximately five pence each. When scanned, the results of each student are automatically recorded on a *Lotus* format spreadsheet. An analysis of the collated results can be generated in terms of facility index, discrimination index and selection frequency for each question in the test. (See Brown and Pendlebury, *op cit*, for explanation of terminology.)

The marking scheme can be configured to take into account the amount of credit or debit to be allocated to correct, incorrect or omitted responses. Student identification numbers are used to register each student and these can be down the master answer card, should a mistake be made, and re-marking of already scanned cards takes placed automatically.

The day after the MCQ test has been completed tutors are given the results, together with notes which highlight the different aspects of the test and its results. The following week, the students attend a remedial session in which they receive the results of the test and the appropriate lecturer covers issues raised by the analysis of the previous week's MCQ paper.

The MCQ tests contribute 15% towards the end-of-semester assessment. From information gathered to date, students seem to like this form of assessment because it gives them regular guidance on how much to learn, and is combined with continuing, accurately targeted remedial sessions.

This system of producing MCQ papers, scanning and analysing the results was developed by Dr Andrew Booth in collaboration with the University of the West Indies Distance Teaching Centres as a solution to their problem of assessing students at sites throughout the Caribbean Islands.

Chapter Four

Self-assessment

Introduction

The term 'self-assessment' refers to a wide range of practices. Self-assessment involves students making judgements and reflecting on what they have learnt. Being able to monitor one's own performance and recognise the need to improve in one area or another is an essential skill in the development of effective learning techniques.

Self-assessment may be used either to judge one's own work, or to help develop it. It is a method of assessment, which should complement other methods, not replace them. It can be used to reflect on actions or on the content of work completed. Self-assessment may be used to discover what has been achieved, how it was achieved, how and why it was done as it was.

Many academics are reluctant to use self-assessment techniques. Doubts persist as to their reliability and validity. However research indicates that self-assessment can be as reliable as other forms of assessment (Falchaikov, Boud, 1989). Ultimately what concerns the majority of academics, however, is the question of who has the power to decide. Do students have a right to judge their achievements? Whether the response is negative or positive to this question, self-assessment still has a purpose in that it helps provide students with the tools to organise their learning effectively and the skill to be self-critical and reflective about what they are doing. It is important to realise that students are unlikely to have any experience of assessing themselves and therefore, 'self-assessment must be practised if it is to be developed' (Boud, 1987).

Self-assessment can be used formatively, summatively or separately from formal assessment, as a task or course requirement, or linked to formal assessment. If used summatively alone, self-assessment involves more than asking students to give a mark to their own work. It is a means of involving them in the process of deciding what is and what is not good work. Consequently one can use any form of checklist, rating schedule or open-ended approach, which a tutor might use for assessing student learning.

When introducing self-assessment it is important to consider the following.

- Have the students done it before?
- How the process and benefits will be explained to them?
- How training and practice will be given to them?
- Is it for summative or formative purposes, or both?
- If there is disagreement, how will it be resolved? Is there room for negotiation?
- What are the learning intentions from this assessment? Do the criteria reflect these?
- Are the criteria for assessment to be negotiated with the students?
- Do the students know the weighting of marks for this assessment?
- Do they understand why they are doing it?
- Is the assessment task designed so that the students are finding out what they have achieved, how they achieved it and why they did it in the way they chose?

Self-assessment and technology

Self-assessment involves students in making decisions about the work they have done; the emphasis is on the student making the decisions. Technology-based assessment involves using an alternative to the human mind to make decisions concerning the assessment. It would be difficult to use the term self-assessment for a system, which involves using a computer to give a mark. However, technology does lend itself to being an excellent tool and facilitator of self-assessment. The majority of technology-based self-assessment is in the form of MCQs, or questions of a similar format.

Self-assessment and MCQs

Many of the CAL and CBL applications developed in a variety of disciplines include an element of self-assessment. This usually involves students being able to measure their understanding of a particular topic through a series of MCQs, which are either embedded in the courseware, or stand alone as a separate package. The package provides the student with a score and in some cases feedback. This is formative assessment, although not strictly self-assessment, because the students are not required to come to any conclusions about the quality of their work.

Technology can be used in several ways to help students to test themselves:

- MCQs which are built into CAL, CBL, and multimedia packages which students complete at various stages to indicate to themselves their progress through computer-based materials;
- MCQs which are available for students to use as a stand-alone package to test their understanding of a previously used CAL, CBL package, lecture or seminar topic;
- Picture-based questions (see Chapter Four), similar to MCQs in their capacity to provide self-tests for students.

Ideally technology-based self-testing should, if it is to be formative, include some feedback. It is more productive for students if, after they have finished the test, a breakdown of their scores is displayed, incorrect responses highlighted and correct answers not only given but explained.

There is some debate as to whether tutors should be able to check that students are undertaking self-tests, if they are intended to be voluntary. The take-up of a self-testing bank of MCQs is an important indicator of students' estimation of its worth, and there are many possible reasons why a high or low take-up rate may be found.

A low take-up rate might occur because students have far too much assessed work to undertake something which is voluntary; the questions could be impossibly difficult or ridiculously easy; there may not be enough accessible computers; students may not be aware the facility exists or they may be technology shy.

A high take-up may indicate that students have little else to do, the questions are easy and boost confidence; or alternatively students could find using the system an enjoyable and beneficial experience.

It is perhaps more important to identify who is using the system rather than how well or badly they are performing. This should not be to establish names and numbers, but to establish whether you have aimed your tests at those who need them. If there is a take-up rate of 65%, the question might be asked - is it the 65% who most need the self-tests? Quite often tutors find that it is the able students who use the system the most, often deterring other students by not freeing computers for their use.

Some practical points to consider when setting up a self-testing system include:

- Do all the students know it exists?
- Do they really know where it is located?
- Do they know how to access the system, or even turn the computer on?
- Is there help, if they have problems?
- Does the system provide good, formative feedback?
- Can the students chose from a variety of topics and levels?
- Can they get a print-out of their results if they want?
- Does the system refer them to references for further information?
- Do you have a method of checking which questions are answered correctly and incorrectly most often?
- Is it possible to obtain information about when the system is used most and by whom?

Technology-based self-assessment facilitators

Self-assessment is concerned with teaching students to judge their own work. It is possible to use technology to facilitate self-assessment by presenting students with a set of criteria against which they rate their work. This method has been developed at The Nottingham Trent University and is described in detail in Case Study 2 overleaf. (8)

Who uses technology-based self-assessment?

Technology-based self-assessment is used in a wide variety of disciplines. There is often the facility of a bank of MCQs available for students to test themselves whenever they wish. A wide range of disciplines, including Mathematics, Statistics, Music, Biology, Engineering, Geography and Psychology, to name just a few, use CAL and CBL materials, which often include some self-testing material integrated or separate from the teaching materials. Increasingly, universities are establishing central computer-based learning units which help guide individual departments in the design and implementation of such materials. In disciplines such as Medicine and Dentistry there is greater use made of picture-based materials to simulate real-life situations and provide students with the opportunity to benefit from specialist knowledge. Picture-based materials could also be used in many other subjects such as Geography, Geology and Environmental Sciences.

Case Study 2 - THESYS, Knowledge Systems Centre, The Nottingham Trent University (8)

The *THESYS Project* developed an expert system package in two parts. The first allows students to assess their own projects and the second is used by project supervisors to identify areas in which students may need more help and to act as a benchmark for assessment. The students use the package when they are about two-thirds of the way through their projects. The introduction explains the keyboard lay-out and ‘Help’ facility. Then follow four sections with a total of 38 questions. Each section tackles a specific area: the first deals with effort, understanding, time management and originality; the second with the quality of the work; the third with the writing-up of the project - content, quality, style, evaluation and background reading; and the fourth with the lay-out. The questions were devised from the Council for National Academic Awards (CNAA) Guidelines, which were designed to be generic, rather than subject-specific. In the composition of these questions, research established where one course might be significantly different from another. Project tutors and course leaders from some of the University's largest and longest established degree courses were consulted. It was found that there were few differences in the ranking given to questions on effort, problem-solving, time-management and presentation. Subsequent usage has confirmed that there is little need for separate versions of the package for each discipline.

The student's package

The questions appear on the screen. There is a scroll bar with opposing statements at each end of the bar. The student must judge his or her position in relation to the statements and record his/her decision by moving a cursor along the scroll bar to the appropriate point. When the student has answered all 38 questions, a final mark is calculated as a percentage. This is displayed alongside a breakdown of degree categories for the University, related to percentage boundaries. A breakdown of the four sections is shown, indicating the best and worst parts of the student's project. A more detailed breakdown of each section is then shown in the form of a histogram. This highlights the effect of weightings on different questions. The student's marks are shown on a swingometer in comparison with their peers and predecessors, and the five best features of the project are highlighted, indicating which areas students should stress in their subsequent oral examination.

The supervisor's package

This is similar to the student's package with different questions in some sections. Supervisors use the package to provide guidelines for report-writing. Over a period of time it has been shown that students' and supervisors' assessments of projects coincide to a large degree. Both the quality of project work and students' self-assessment skills have improved since the system became widely used. A statistical analysis, undertaken to compare traditionally marked and *THESYS*-generated results, revealed that there was no significant variation.

The software package was written in the *Leonardo* expert shell system, for which licensing has proved to be expensive, and customisation limiting. *THESYS* runs on IBM compatibles with VGA/EGA graphics and a minimum of 640K memory. The project to create and validate the software was funded mostly by the Employment Department.

Chapter Five

Coursework and Examinations

Introduction

The word 'coursework' has a range of meanings. Coursework is defined here as work, which must be completed by a student. It is a formative assessment contributing to a final mark, and is not completed under examination conditions. Both examinations and coursework may include practical work, essays, problem-solving exercises and some more discipline-specific tasks, for example language composition and computer-programme writing.

Essays

Essays are commonly used in the Arts and Social Sciences. They can require students to combine knowledge, understanding and skills, to identify a problem, create a structure and present an argument. Essays in Engineering, Medicine and Science often resemble factual accounts, although questions can be set which require students to identify connections between concepts, approaches and methods of analysis. In most essays there are no right or wrong answers; it is the selection of appropriate material and presentation of an argument which determines the grade. Essay questions often seem easy to set but are difficult to mark objectively. Brown and Pendlebury (1992) give helpful details of essay forms of assessment.

There is much literature to suggest that the marking of essays is unreliable and assessment of the same piece of work, by different examiners, produces marks, which differ considerably. Research by Bell (1980) and Rowntree (1987), indicates that variety in grades between examiners is common, and that there is little agreement even between 'experts' as to what makes a 'good' answer. This makes the task of producing a 'good' essay difficult for the student.

It is the lack of identified objectives and criteria, which tends to hinder the reliable grading of essays. The development of objectives and criteria will help to minimise bias and inconsistencies. For further discussion of developing assessment criteria see Brown and Pendlebury (1992).

Essays and technology

There is presently little if any use of technology to assess students' essays. This arises because the nature of using technology to assess is such that it must first be known what is to be assessed. To use computer software or other technology it is necessary first to establish what qualities or responses the system is to identify and grade. It is often difficult to identify in advance what students may write in a particular essay, and to decide the full range of possibilities to be expected. In the Arts there is also the added complication of there being an infinite number of possible answers, all with their own merits. It is also possible that students may develop an unexpected interpretative framework. Essay marking by computers is not yet possible and some would say, not desirable. However it is possible to use many packages for spell-checks, grammatical consistencies and measures of readability. These procedures may of course be applied to essay questions as well as essay answers. It is also possible to mark short-answer questions in which key words and unique statements are required. As yet there does not appear to be any major development along these lines.

There are also however, a number of ways in which technology can be used as a tool of assessment. These are investigated in Chapter Six.

Practical work

Practical work includes laboratory work, fieldwork or clinical experiments; it has long been used in Medicine, Engineering, Social Sciences, Education and the Sciences. Practical work is perceived as encouraging students to learn more actively because they are engaged in performing a task, rather than listening or reading. It is however an expensive method of assessment, and evidence suggests that laboratory work focuses mainly upon manual and observation skills (Hegarty-Hazel, 1986).

Practical work can help students to develop problem-solving skills, improve their observation skills, improve their understanding of methods of experimentation and help them to develop professional attitudes (Brewer, 1985). To be effective, practical work must be seen to be relevant by students, and must result in good feedback. The assessment of practical work has been found to be unreliable. Wilson *et al*, (1969), found a 25% variation between sets of demonstrators who had marked the same laboratory notebooks. In practical work the objectives of the assessment must match the course and experiment objectives for it to be a useful and efficient assessment.

Practical work can be assessed by using only the write-up, which the students complete once the session is over, or by watching and noting the procedures being undertaken. Some of the problems associated with the assessment of practical work may stem from the use of untrained postgraduate demonstrators to assess the work of undergraduates.

Bliss and Ogborn (1977) identified that demonstrators were uncertain of their roles and students reported wide variations in quality of help and assessment by markers. Daines (1986) reported that many students experience unfair marking, negative criticism and poor explanation of errors from laboratory demonstrators. Daines identifies that students valued ‘fair marking, constructive criticism, clear explanations, useful written comments and listening to students’ questions’.

When the objectives of the practical work are established, it is important to clarify what is to be assessed in each experiment. In some practical sessions concentration may be on the description of a procedure and collection of results; in others the focus might be on the analysis of results and their interpretation. It is essential that both lecturer and students understand and agree the objectives of the experiment. Once again more detail on assessment in practical work can be found in *Brown and Pendlebury* (1992).

Practical work and technology

There are several ways in which technology is used to assess practical work.

- Computerised/optically scanned MCQs can be used to test the understanding of laboratory procedures, and the interpretation of results.
- Computer-marked practicals - in disciplines where the main objective of practical work is for students to be able to record accurate measurements and results, and make correct calculations using the correct procedures. Computer software can be used to check calculations and accuracy. An example of computer-marked practical work is given in the case study section of this chapter.
- Picture-based computer programmes - these involve the use of images and text presented on screen. These programmes can be used to enhance practical work by presenting information , which would otherwise be unavailable or difficult to create or obtain in a laboratory or clinical situation. Students may be asked to label objects on the screen, make calculations, answer questions and indicate a physical action as a result of a specific outcome. The computer programme can give them feedback and can record their responses. Programmes such as these are often used in a tutorial or self-teaching environment and are most widely

used in Medicine and Dentistry. Examples of such programmes include the *Orthopaedic Surgery Tutorial Package at Queen's Belfast* (9), and *the OSCAL (10), Oral Surgery Programme* at Leeds Dental Institute.

In addition, a standard checklist could be used and once completed, optically scanned and analysed. This would provide rapid feedback to students and save administrative time.

There is great potential for using technology as a tool to aid the assessment of practical work. Such uses are detailed in Chapter Six.

Problem-solving activities

Problem-solving is a term which may encompass a range of activities. It could refer to solving a problem where both method and solution are known, where method only is known or where neither is known. As the complexity of the problem increases, values are likely to intrude and this makes assessment more difficult. Problem-solving is a method of assessment used widely in Mathematics, Engineering, and Computing. However it is possible to use problem-solving activities in a range of other disciplines such as Geography, History, Economics and Philosophy.

It is important when assessing problem-solving exercises formatively to provide constructive, meaningful feedback, which helps students to understand how and why they went astray, if indeed they did. The assessment of the task does not necessarily have to be based on the correctness of the final answer - it is the process, which is being assessed, probably more than the answer. It may also be the case that for some problems there is no correct answer, only superior ones.

Assessment of problem-solving activities usually takes the form of a written examination. However, this could prevent students from learning from each other. Learning from peers may be a most productive way for students to learn, considering that many tutors are so adept at solving problems that they have difficulty in understanding those who do not have a good grasp of problem-solving skills.

There are a variety of methods by which to assess problem-solving skills. Some examples are:

- Incomplete solutions. Part solutions are provided to defined problems containing common errors. Students are asked to complete or correct.
- Presentations. Students (individually or in groups) are set a practically-based problem, about which they have to write a report and give a presentation to their peers.
- MCQs. MCQs are used in coursework to test students' grasp of particular concepts, and their ability to identify correct solutions and arguments.

- Peer support. Each student is given a short problem and a limited time in which to investigate it. Then the students discuss the problems in pairs and help each other develop solutions. They then assess their contributions to each other's problems.

Problem-solving and technology

The extent to which technology can be used to assess problem-solving exercises depends to some degree on the nature of the problem, how definable its steps are and whether there is a single correct answer. It is easier to use technology to assess, if the problem-solving exercise is coursework rather than an examination. Much of the work in this area is confined at the moment to the assessment of mathematical problems. However, the example given below indicates that there is potential for the adaptation or manipulation of software used to predict outcomes and to assess the impact of various factors in industry.

The School of Built Environment at Anglia Polytechnic University uses a software package called *Energy Targeter* (11) to assess students' design and problem-solving skills. It is a sister package of *Energy Evaluator* used by the National Energy Foundation to conduct rated energy audits. The package was modified for student use. Students enter specific details about the design of a dwelling and have a number of variables, which they must adjust to attain a specified energy target level. The programme assesses the energy efficiency of the proposed dwelling. It warns students if they are breaking building regulations and provides them with the resulting energy target. Students can go back into the programme and alter variables in order to attain a better energy target. This software is used in formal, formative assessment of students' work. A similar programme exists to assess the daylight efficiency of dwellings designed by students.

There are a number of programmes in use designed to assess mathematical skills. The *CALM Project* (13), a computerised tutorial system, uses in some cases problem-solving exercises to enhance the teaching of calculus, and is described in the case study section of this chapter. There also exist a number of other examples of computerised learning and assessment in Mathematics, such as the *CALMAT system* (14) at Glasgow Caledonian University, a collection of 50 CAL units including randomly-generated post-tests and progressive exercises. An evaluation of the *CALMAT* CAL units was undertaken by Coventry University and results are detailed in the Quarterly Newsletter, February 1993, published by CTI Centre for Mathematics and Statistics, Birmingham University.

Some of the existing technological methods of assessing problem-solving exercises are used for formative assessment only and do not contribute towards a final assessment. Others are required to be completed before progression - the *CALMAT* system is used in the University's summer school for those students with a conditional pass who are therefore required to complete prescribed units before acceptance onto the following year. There is however scope for the use of some of these packages to conduct summative assessment. In situations where each student could be given a unique problem which could be assessed by the same software, or where tests are randomly generated, the perennial problems of plagiarism could be overcome.

Subject-specific coursework assessment

There are some types of assessment, which are on the whole subject-specific. In disciplines such as Medicine and Computing, students are asked to complete assessment tasks, which are individual to their subject, such as design software, or examining and diagnosing a patient. That is not to say that the skills and understanding being assessed are also unique to their subject. The writing of a computer programme could be considered to be similar to writing an essay, in that it requires logical thought, style, fluency and expression; or similar to a problem-solving exercise, in that it requires the student to work through a series of steps, creating a functional structure.

In the discipline of Computing there is perhaps not surprisingly more development of automated assessment than elsewhere. The assessment of programmes written by students can be time-consuming and sometimes tedious, and the third case study at the end of this chapter describes a software system developed at the University of Nottingham to mark programming skills automatically. Other institutions are developing similar programmes: the School of Computer Studies at the University of Leeds has a small in-house submission and testing system (15); the Computing Department at The Nottingham Trent University is developing a system to automate and speed up the evaluation and feedback process of programming assignments (16).

Examinations

Examinations are usually compulsory assessment tasks performed under formal, time-constrained conditions. They could include any of the above methods of assessment. There has been criticism of examinations in recent years, despite their value. Like other assessment methods they must be carefully thought out:

The reasons for examinations for many are that they:

- are more effective than coursework in preventing collusion and plagiarism;
- are well-known, and tried and tested;
- are relatively cheap and efficient;
- are required (in some disciplines) by professional bodies;
- may assess what cannot be assessed by coursework;
- motivate students to learn.

There is much debate about the worthiness of the reasons listed above. The last two points especially come in for criticism. It is often not clear as to what exactly examinations assess which coursework does not. Although examinations can motivate students to learn, it is the type of learning, which they encourage, which is often under question. Examinations are criticised for encouraging surface learning, with students memorising vast amounts of

knowledge, reproducing it in an artificial situation, and consequently rapidly forgetting it. However, it should be noted that surface learning may well be a valid skill, and in some professions the assimilation of large amounts of knowledge to be used for a short period is required.

The practice of allowing students a wide range of choice of questions in examination papers also creates questions of *reliability* and *validity*. The *reliability* of an assessment can be measured in several ways. The main measure of reliability is measures of agreement between assessors. The use of agreed criteria, second markers, and processes of moderation are ways of minimising errors. Reliability can also be enhanced by increasing the sample of assessments. This however may result in overburdening students and tutors. A second measure of reliability is the internal consistency of the assessment, this measure is used to refine MCQs. This almost eliminates the use of judgement by the tutor, as interpretation and correctness are decided by the person who constructs the test.

Validity is often described as how closely what is intended to be measured matches what is measured. If objectives closely match the assessment tasks, internal validity is likely to be high. Validity can be measured in terms of appropriateness; concurrency; predictability; and criterion validity. Appropriateness measures whether the results of the assessment meet the objectives of the course. It draws on the fact that objectives must be clearly expressed and are measurable.

Concurrency measures whether the results of the assessment match those by a similar assessment task. This is difficult to achieve in practice, and comparisons are usually made between written work and coursework. Although correlations between the two are usually high, they are not perfect.

Predictability measures whether the results of the assessment predict future performance accurately. This relates selection procedures to subsequent assessments and destinations of graduates.

Criterion validity measures whether the results of the assessment can be used to discriminate between those who have taken a course and those who have not.

Examinations should provide students with the impetus to review their work, fit it all together in a coherent structure and make connections, which they may otherwise have missed. Factual questions, which relate directly to lecture material, usually suggest to students that the memorisation of lecture material is all they need. Finally, examinations do not necessarily have to be based on 'unseen' questions, and can be presented as 'seen'

examinations, and 'open-book' examinations. Examinations can be used formatively if meaningful feedback is provided, rather than just a mark. For more information about the application of different types of examination see *Measure for Measure: An Assessment Tool Kit*, Lancaster University, Unit for Innovation in Higher Education.

Examinations and technology

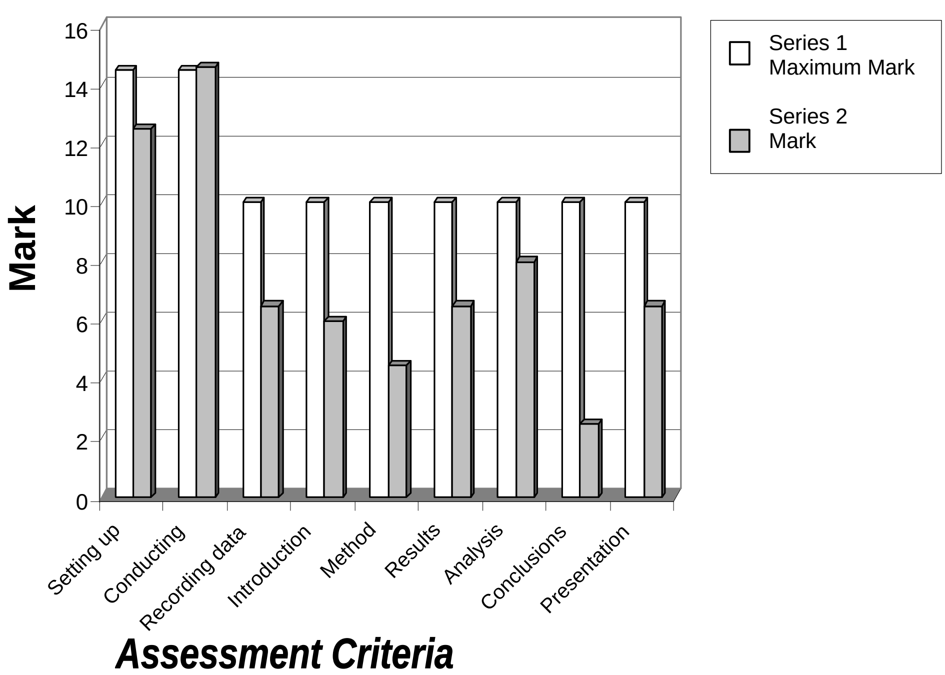
Many institutions use technology-based processes to collate and analyse examination results. University-wide networked systems can be used to handle large data sets, compile, analyse and print out examination results, but are not common. In formal examinations technology is mainly used as a tool of assessment and this is discussed in more detail in Chapter Five. There are few examples of formal examinations in which students sit before a computer, enter their answers, which are then automatically marked, recorded and printed out. Appendix 2 outlines a range of computerised assessment systems, ranging from a paper-based to a fully automated system. A fully automated system could be a fast, fault-free, efficient method of assessing large numbers of students. It is however, if used regularly, dependent on the availability of resources, ability of students, and would have to be carefully policed if the security of the system were to be maintained.

The final case study in this chapter describes an examination taken on computer in the Department of Italian at the University of Hull (18).

Technology and group performance

Technology has a role in assessing the overall performance of a group of students. Overall class performance can be assessed either through the completion of a checklist entering group totals for each section, or through spreadsheet software, which can analyse class performance and provide graphic representation for each criterion (see Figure 1). This provides tutors with information to assist in their subsequent teaching. Using standard deviations and ranges can also be used to explore in more detail aspects of overall class performance. Figure 1 shows a simple graphical representation of class mean performance in a practical class. This information may be provided to the class and can be used to assist in the development of teaching.

Figure 1 - Practical Class Results



Case Study 3 - Department of Geology and Physical Sciences, Oxford Brookes University (12)

This programme was designed to:

- help assess the level of practical skill shown by a student as indicated by the agreement of her/his results with averages from the entire practical group;
- allow students to check calculations based on their own experimental data.

On completing the practical component of a laboratory session, a student goes to a computer, which is running the data-input programme. The computer prompts for the student's name and then for the experimental results. These are written to disk. The computer than asks the student for answers to quantitative questions based on their own results. The computer records which questions the student gave correct answers to and then informs the student which questions were correctly and which were incorrectly answered. The student is allowed repeated attempts at questions, which were incorrectly answered, with the computer maintaining a note of how many attempts were made at each question. At the end of the session, a marking programme gives marks to individuals based on the closeness of their mark to the class mean, and the number of attempts required to answer the quantitative questions correctly. The class list and assigned marks are printed out.

The advantage for students is that they can check their calculations immediately without waiting for laboratory reports to be marked. This enables them to identify any calculations for which they need additional help from staff. The ease of calculating class mean scores and assigning some marks on the basis of how close observations are to these mean values encourages students to improve their manipulative skills.

Staff can use the programme to ensure that students complete their calculations prior to leaving the laboratory session. The automatic check on numerical aspects releases staff time for assessing more descriptive elements of laboratory reports. An authoring programme allows a member of staff to modify the prompts given to students, and to enter the BASIC code required for the computer to calculate answers from student experimental observations.
(Gibbs, Habeshaw, Habeshaw (1988) 53 Interesting Ways to Assess your Students)

Case Study 4 - CALM Project, Heriot-Watt University (13)

The *Computer-Aided Learning in Mathematics (CALM)* project is a computerised tutorial system for teaching calculus. It began in 1985 under the Computers in Teaching Initiative (CTI) and produced 25 units of courseware covering various aspects of calculus. *CALM* is used to teach Mathematics to engineers and scientists and includes problem-solving exercises, which aim to bring the applications of Mathematics to life.

The units allow students complete control over their route through the tutorial. Although the summative assessment for each unit is conducted in a conventional manner, each unit has a test section, which allows formative and diagnostic assessment. The test section of each unit has three levels and offers different types of help at each level. It enables students to identify their strengths and weaknesses, and improves their ability to assess themselves.

Students' marks are recorded and their progress is monitored. This helps to identify students who are having difficulty, and tutors can then send them appropriate help.

The *CALM* assessment shell is presently used at other higher education institutions, including Brunel University, Southampton University, UMIST and Essex University.

Case Study 5 - CEILIDH, Computer Science, University of Nottingham (17)

CEILIDH (Computer Environment for Interactive Learning in Diverse Habitats) is a system which was developed to aid the teaching of programming languages and includes as a major part of the system an on-line coursework submission and auto-marking facility. The system also has exercises, notes and solutions available on-line.

CEILIDH has student, staff and teacher menus. The student menu allows students access to:

- general course information such as deadlines for submission of work;
- specific information about the coursework they must write;
- outlines of programs; and
- model solutions.

It also allows them to ask the system to mark a program and to submit coursework through the system. There is a 'Help' facility and students can comment on particular exercises or on the system in general.

The staff menu includes all the above, and allows staff to list details of work submitted and names of those who have not submitted work, to list marks awarded for each particular exercise and to check attendance at laboratory sessions. The teachers' menu additionally allows teachers to set new coursework, amend existing coursework, distribute marks, discover the strengths and weaknesses of the whole class, run a plagiarism test and browse through the comments file. Individual graphs can be produced, with each student's marks shown in comparison with his or her peers.

The system assesses each student's program for style, layout, design and format. A mark history is kept for each student. Students are allowed to use the program to assess their work, and to go back to their program and improve it as many times as they like before submission. This aspect of the system supports the view that, if students can immediately and repeatedly re-mark their programs, they will work towards a higher level of quality. This represents a key innovation in assessment technique.

However, this aspect of the system has created some problems with students, who have in effect become addicted to the system, and repeatedly re-write their programs to gain one or two additional marks.

On the other hand the system has resulted in improvements in the quality of student feedback; the amount of administrative work has been reduced; and staff time has been saved through the reduction in marking hours.

Case Study 6 - Department of Italian, University of Hull [18]

The following case study describes the use of a computer test for beginners in Italian.
'Computer testing for beginners in Italian

Our beginners' Italian module (for about 80 students) is formally assessed at the end of the year. Last session the written test taken before Christmas to check on progress was transferred to computer. Student feedback was good and staff time for marking 80 scripts was saved.

We used the Italian version, 'Che aggiungo?', of the Wida Software and Eurocentres authoring program, Gapmaster. There are two modes, 'Modo esplorativo' (Explore mode) and 'Modo accertamento' (Examination mode). We used 'Modo accertamento' which gives a percentage mark at the end of the exercise. The computer test closely resembled a written test. Instructions and an example for each exercise appeared on screen, there is an introduction and help screen in the program, but we considered them more of a hindrance than a help for this particular kind of test.

We chose the Italian version to introduce our beginners to computer language in Italian. We also hoped to encourage more advanced students to program further exercises for use in the Department as part of their computer projects for language assessment. In retrospect, the Italian version turned out not to have been such a good idea. The English version has been updated to Gapmaster , but Wida Software Ltd and Eurocentres are not considering a new Italian version at the moment, due to lack of demand.

There are some new features on Gapmaster2, which would be useful for our test. There is now a function keys option with a list of accent characters. For last year's test we provided our students with such a list. In the old version there is a maximum length per exercise of 10 pages and a maximum of 12 lines per page. This meant the test had to be split into 2 parts. In the new version exercises can have a gapped text of no pre-set length with up to 120 gaps, so our test would no longer have to be divided. When a student finished the first part, the invigilator collected the mark from the student's screen. The new Gapmaster enables students to save their work; record their score and both student and teacher can review the answers. In view of these extras on the updated version we will probably move our test over to the English version.

Both versions are straightforward to use. I had some problems, but these were mainly due to the fact that I had not read the manual properly. When you leave your gapped text, the program informs you if there is anything wrong.

The test itself included mostly function gaps (definite and indefinite articles, articulated prepositions, etc) with usually only one correct filler. In one section the test was partly semantic as well. For example, for the sentence 'Quell'uomo ha tre bei figli' (That man has three beautiful children), the correct form of the adjective 'bei' had to be provided, and the student was given the English word 'beautiful' at the start of the sentence.

Students are prepared for the test by a weekly computer class, also introduced last session for our beginners. The rest of the beginners' Italian module is made up of grammar-based classes with a communicative approach, a class in the Language Centre using a variety of media and a role-play class. In the computer class students do drills using the Leeds University program, LUISA, to consolidate their learning of grammatical forms.

Since I began teaching some ten years ago, our beginners' language course has moved steadily towards a more communicative approach, but we have not been entirely satisfied with the results, since written accuracy suffered at the expense of confident communication. We have found in the past that the speed of learning we demand can result in disaster for the weaker or lazier student who accumulates a mishmash of half-remembered grammar and vocabulary. For this reason it is vital to check student performance at Christmas before they proceed any further. The computer class and test has made students more conscious of written accuracy and LUISA is popular with the students because they can work on it at their own speed and do extra hours in their own time.

The test took place during class time, in a one-hour period. Students with a mark under 50% were asked, as in the past, to repeat the test in January. We noticed improved results last year, only about 15% of students had to resit, and none among those students doing Italian as a major part of their degree. We have found Gapmaster very suitable for this level of study. Formal assessment at the end of the year is naturally more sophisticated, although there is still a grammar section for which we will use the program next year. Text already typed can be transferred to the program, which means interim tests from the past can be made available to students, or we can create further exercises for use with LUISA.'

Dr Pamela Williams

Chapter Six

Technology:
a Tool for Assessing Students' Work

Introduction

As mentioned earlier in this paper, technology can be used as a tool of assessment - it *aids* assessment procedures but does *not replace* the cognitive processes which a tutor must undergo in order to arrive at a mark for a piece of work.

This section outlines the possible uses of a range of different technologies, and gives some examples from five disciplines. Many tutors keep records of their students' marks on a spreadsheet, or allow/require students to submit coursework either word-processed or on disk. In some higher education institutions tutors and students exchange assessed work, marks and feedback via Email. It is also possible to use computer software to automate the process of giving feedback to students about their coursework - to the extent that a touch of a button results in a print-out which contains comments and suggestions relating to a particular essay.

This section intends to highlight some of the more interesting and innovative methods currently being used.

Video

Through video it is possible to give students the opportunity to experience situations which would otherwise be unavailable to them. In some disciplines students may be required to make a video as part of their assessed project/group work.

Video can be used:

- to enhance lecture material - some topics are better described visually, eg the science of building demolition;
- illustrate a particular concept or theory;
- as a focus for group activity;
- to encourage media literacy;

- to promote communication skills;
- to assess a different range of skills.

Video could be used to present a problem in a lively, visual way. To enhance problem-solving exercises, solutions could also be shown. It could be used to help students to identify problems and develop solutions. In some disciplines video is used to record practical and clinical work, create role-playing situations, which can be played back, discussed and analysed.

Video is presently used in disciplines such as Psychology, Law, Medicine, Social Work and Management. The following describes the use of a video case-study examination.

Case Study 7 - Financial Management and Accounting, University of Bradford Management Centre (19)

The use of a video case study examination was first introduced in this Centre in 1992. The idea behind using a video case study was that it would bring the subject (Accounting and Finance) to life and would give students an opportunity to use knowledge from both their first and second year courses. The examination question was 25% of the total examination marks. There were 106 second year Business and Management undergraduates involved.

As part of normal lecture sessions a series of three videos was shown. The students were not at this point aware that a video case study would be part of their examination. A month or so later the students were informed that the examination would include a video case study. Repeat showings of the videos were given at that stage and again, prompted by student demand, following the Easter vacation.

The question was one of four, two of which had to be answered. It proved to be the most popular question on the paper. The average mark given for the case study question was higher than for the other two essay questions on the paper.

It was concluded that as a result of the video case study:

- Students' motivation increased. Attendance at lectures increased at a time when it usually decreased, and students were making an effort to understand and digest the content of the video case.
- A wider area of knowledge was covered by students, because the video case did not require them to treat financial information in isolation from whatever else was going on around.
- Students' comprehension and appraisal skills improved.
- It was a suitable method of teaching a large group and could have been used to teach a group of several hundred.
- Students benefited from forming their own tutorial and discussion groups to exploit the video case material.

Computers and examinations

There are a number of subject areas in which computers are used by students in examinations. There can be problems of resources and some students may lack keyboard skills. This area may see particular development in the next few years.

In some disciplines computer software is used regularly in coursework, as part of the teaching material, or is often required by employers of students seeking vocational/professional qualifications. It may therefore be reasonable to use computers in an examination, and allow students to demonstrate their computer literacy.

The combination of examinations taken at computer terminals, marked by computer software with results automatically recorded, processed and compiled, and resulting in degree classifications, without the paper ever touching a tutor's hands is technically feasible. It is, however, a long way in the future, and perhaps not even desirable. Using a computer rather than paper and pencil does, however, in certain disciplines provide a method for testing concepts and skills, which may be difficult to test otherwise.

Students use computers in examinations in several subjects such as the use of statistics packages by second year Psychology students at the University of Newcastle upon Tyne in examinations. Foreign language departments also use software packages designed to test vocabulary and grammar, such as the Fleece *'Teach-yourself-German' package* and *LUISA*, an Italian teaching package developed at the University of Leeds.

Case Study 8 - Management Information Systems, University College, Dublin (21)

The department conducts practical examinations, using a variety of software which students have used in their courses. Three subject areas are tested in this manner: computer applications; operations research; and statistics. Conceptual knowledge of the material is required as well as knowledge of the software package.

The examinations range from one to three hours in length and are taken on a PC. In 1992/3 about 1800 students sat these examinations in 45 sittings. The examinations included:

- a 2.5 hour spreadsheet and word-processing examination;
- a 1.5 hour database examination;
- a 1 hour statistics examination;
- a 3 hour Management Science examination, where use of a PC is optional.

Each sitting has 30 to 50 students and there may be as many as 500 students taking any one type of examination in 10 to 12 different sittings. The exercises given to the students are slightly different in each sitting - in spreadsheet questions the numbers used are different and there are some larger variations in the question type used. Question papers are numbered, and students must write their name on the paper and hand it in when leaving the room.

Part of the marking of papers takes place during the examination. Examiners mark graphs or screen designs which students have created. Students save their work on disk and tutors mark it after the examination. Students' work is stored on a network server and students have special examination 'log-in' codes which allow them access only to those files and programs which they are permitted to use in the examination.

There are a number of problems:

- **Time-tabling is difficult.** Groups are allocated randomly or alphabetically, and tests are during term-time when students have other academic commitments. An experiment to allow students to enter for the examination which suited them best, caused problems because resulting groups had significant differences in ability.
- **Resourcing and organisation**. Only 30 to 50 PCs are available which restricts the size of the sitting. Because there are multiple groups, multiple sittings are needed.
- **Nature of examinations**. This creates difficulties in ensuring that the last examination is of equal difficulty to the first. Despite efforts in collecting papers along the way sufficient information may leak out to make the last examination easier than the first.
- **Minor technical problems.** These arise generally or in one machine and a student may exaggerate the significance of this to get extra marks. Established procedures cannot deal with this.
- **Staffing.** Only experienced tutors can be used as invigilators. Staff must be able to cope with technical problems and a wider variety of questions from students than in a conventional examination.
- **Inexperienced students**. A few students are not skilled enough to save their work, and gain a zero mark as a result.

However the use of computers as a tool for assessing students' work is undoubtedly worthy of further exploration.

Multimedia, hypermedia and hypertext applications

All these applications have similar capabilities for aiding the process of assessment. The ability to present large amounts of information in a range of different media provides good opportunities to assess students' work in interesting and innovative ways. Applications such as those mentioned above have the potential to present a student with information from a variety of sources simultaneously.

A multimedia package, which presents text, images, video-clips and graphs about a subject can replace lectures and provide stimulation for animated assessment tasks. Students may be requested to work through a hypermedia or hypertext package, which guides them on a particular route or requires them to choose a particular area in which to specialise. The package could offer them a number of different assessment tasks, provide references for them to follow up and encourage active, participative learning. As more and more multimedia, hypermedia and hypertext applications are developed and used as courseware, there will be the need to consider the way in which material taught in this manner is assessed.

In the Department of Italian at the University of Hull a Guide Hypertext project on Petrarch was substituted by two fourth-year students for an assessed essay (22). The project included the annotation of a poem, a commentary on the text, a metrical and syntactical analysis, a glossary of rhetorical terms, a reference section on main themes and a bibliography. It is now a resource from which other students can learn. The following case study illustrates how a multimedia application is used to teach and provide an aid to assessment in Anatomy, Physiology and Histology.

Case Study 9 - Biological Sciences, St Andrews University [(23)]

MacCycle is a multimedia application used to teach second-year medical students about the menstrual cycle. The software replaces two lectures and between three and five hours of practicals. It integrates the teaching of Anatomy, Physiology and Histology.

The software consists of text, graphics, still pictures and video-clips. Students can look at photomicrographs of histological specimens at different magnifications, see dynamic changes in hormone levels and study explanatory diagrams.

The application takes about two hours to work through completely. Once students have done this they are asked to write an essay electronically. The essay may include resources taken from the original programme. Images and graphs, for example, can be used to illustrate certain points.

There is a large bank of resources and students must be able to sift through the large quantity of information and identify important points. The essay is saved on disk and given to the tutor who prints it out to mark it. Tutor time is saved because all essays are legible, and the programme limits the number of pages which an essay can contain. Students are encouraged to work through the programme, read and plan their essay and return to write it later. The majority of students have already done some computing and do not find the software difficult to use.

The programme was written in *Hypercard*, using *Quicktime*. Slides were optically scanned, and a motion video adapter, frame-grabber and flatbed scanner used to insert video-clips. *MacCycle* contains 42 photomicrographs, 30 annotated diagrams, 4 video-clips and 2 graphs of hormone levels.

It is intended that in future students will be requested to create a module from the information in the application as part of a small group or individual project for assessment. The department is presently developing *MacGut* - a similar application designed to teach the histology of the gastro-intestinal tract.

The Department of Mathematics and Computational Sciences have developed a similar application, also using *Hypercard*. *Mathematical MacTutor* is a computerised laboratory for conducting mathematical experiments. It includes animated displays, MCQs, demonstrations and historical information.

Chapter Seven

The Management of Assessment

Introduction

As the quality of assessment in higher education comes under increasing scrutiny by both external and internal quality systems there will be a need for demonstrable efficiency, accuracy, objectivity and quality assurance in the management of assessment.

Modularisation may result in a larger number of assessments being required and adds further to the pressure to produce results efficiently and accurately. It also requires that the assembly of assessment information takes place several times during an academic year. Students are rightly given more assessment information than previously and universities might, as a result, be under greater risk of litigation. Students can now examine individual module marks and have the means and the right to know how these translate into a degree classification. Questions facing those developing modularised programmes may be how much, and what type of assessment information is stored and for how long. In a modular system students have greater control over the duration of a degree programme. S/he may undertake it in several different institutions, which will require inter-institutional transfer of assessment information.

With growing student numbers and the consequent increase in the number of assignments to be assessed, the probability of error is increased. Greater student numbers without parallel increases in resources emphasises the need to develop assessment management strategies which save not only staff time but also other resources, and which result in a speedy, accurate organisation of assessment information.

Managing assessment by using technology

It is in this area of assessment that technology-based management systems have a tangible and productive role to play. Paper-based assessment management systems can be wasteful in several ways:

- **Staff time.** Academics may spend considerable time compiling lists of class marks, aggregating marks and checking for errors. Administrative and clerical time will be spent in: (a) compiling lists to be sent to other Schools/departments and to the Registry, where the information may have to be re-compiled and checked, entered on to central records and returned for confirmation;

(b) checking that information provided is correct;
(c) printing out letters, address labels to send students' results and
(d) answering queries from staff and students.

- **Inaccuracies.** These can have far-reaching results, depending on their extent. Students increasingly have access to assessment information which was previously regarded as confidential and can query module marks and their components. The identification of an error may result in immediate adjustment of marks and alteration of records, or may lead to an appeal.

- **Resources.** There are resources other than time associated with assessment management systems. Some systems may use large amounts of paper and require records to be stored, which involves the utilisation of space, a resource often in short supply.

There are great savings to be made through the implementation of a technology-based assessment management system. Although initial cost may appear to be high, it is becoming less expensive as the hardware needed becomes cheaper, and in some institutions the hardware may already be on-site and not being fully utilised. There are several advantages to a system, which allows the automatic processing, administration and management of assessment information.

The optical scanning of a student's marks can initiate an automated process which results in marks being aggregated, analysed, partially interpreted, entered on to a student's file, distributed to staff and departments, and printed out as a hard copy for student feedback. A system such as this would: (a) allow efficiency gains in academic and administrative staff time; (b) save resources such as storage space and paper; and (c) may result in a higher degree of accuracy than a paper-based system.

A fully automated assessment system would require all aspects of the assessment system to be technology-based - from the actual assessment which the student completes to the processing and administration of the marks, and the overall management of assessment information. Appendix 1 outlines various degrees to which assessment systems can be technology-based.

Higher education institutions vary greatly in the degree to which their assessment management systems are automated. The Open University has automated its assessment procedures to a high level; this is examined further in the case study section of this chapter.

There are a variety of methods for implementing technology-based assessment management systems. It may be possible to complement existing systems, rather than implementing a new system altogether. At Napier University (24) a computerised management system has been created for their modular courses. A database has been set up containing:

- details of course title and content;
- module aims and objectives:
- which students can take each module;
- which modules are contained in different degrees;
- the number of hours of lectures, tutorials, seminars, direct and independent learning each module requires;
- the amount and type of assessment for each module.

Another database contains students' personal and academic records. The databases are linked and students have access to the module descriptor database. The module descriptor gives departments the opportunity to compare types and numbers of assessment between courses, as well as the assessment load for each module. Once work has been assessed the mark is entered on to the student's record at faculty or departmental level, as a means of recording and transferring information. A system such as this may be extended to incorporate the aggregation and analysis of assessment information.

It may also be possible to re-organise and adapt the use of existing technology, to enhance the efficiency of assessment management. The case study section also gives an example of the adaptation of an existing system to incorporate the management of assessment information, as well as an example of how additional technology can be used to enhance the management of assessment.

Design of assessment management systems

Whilst most would welcome the efficiency gains which can be made with the aid of a technology-based assessment management system, there should be careful consideration exercised in the design and implementation of such systems.

- **Lack of compatibility.** This is probably the greatest hindrance to a fully automated assessment system. If Schools/departments do not have computer systems, which are compatible with each other and with the central administration, it may be difficult to implement a system, which results in efficiency gains on an institution-wide basis. It may be unreasonable to expect institutions to invest in new hardware, when they believe their existing hardware to be sufficient, although not compatible for cross-departmental purposes.

However the gains to be made from an institution-wide, compatible computer system are great and this should be borne in mind when purchasing new hard and software.

- **Flexibility.** This is related to the compatibility of a system. For a technology-based assessment management system to result in real efficiency gains it must be relatively easy to adapt to changes in assessment procedures and to departmental differences. The system should be sophisticated enough to deal with possible system failures, should incorporate a checking procedure to guard against human error and should be secure against would-be hackers.

- **Ease of use.** Systems, which are user-friendly, comprehensive, and include a help-facility will probably reduce the number of training hours needed for staff operating the system. It is helpful to identify who will be using and operating the system most frequently and what systems they are familiar with. The system may be designed with differing levels of access. It may also be possible to retrieve and compile information, which is otherwise inaccessible within a paper-based system. Academic, administrative and clerical staff may well have different requirements of the system and these will have to be identified and incorporated in the planning and design of any assessment management system.

Case Study 10 - The Open University (25)

The Open University has a highly automated assessment system, which has been operating for many years. The system was developed to handle the large amount of information generated by the assessment of large numbers of students across the world.

In the Open University there are two main forms of assessment: continuous assessment and examinations.

Tutor-marked assignments (TMAs) are mainly essay-type or short-answer questions, although some courses require extended essays or projects. Each TMA requires an accompanying pro-forma to be filled in with comments and marks. The marks are entered on to the system and automatically updated in the student's record. In 1992 over 611,000 TMAs were processed.

Computer-marked assignments (CMAs) comprise a series of questions for which students select the answers on an optically marked reader form. The scanning of these forms automatically records a score on the student's record, and generates a personalised feedback letter. In 1992 180,000 assignments were marked in this way.

Scores from the end of course examination is also held on a student's record. The assessment system allows for the statistical adjustment on examination scores at three levels; by scriptmarker, by question, and by overall examination score. Prior to the awarding process the system conflates a student's continuous assessment score with the examination score, to produce a provisional overall course score. This may be confirmed or changed by the Examination and Assessment Board prior to the ratification of results.
Historical assessment data can be obtained from the system.

Case Study 11 - Leeds Metropolitan University, Student Information System (26)

At Leeds Metropolitan University a university-wide, networked, computerised Student Information System (SIS) was developed and implemented to provide central and standardised student information. The system holds details of the academic provision needed for each course/module and student's personal details.

The SIS holds information about formal assessment processes. At present, only the final outcomes of the various degree levels are held on the system. Within 48 hours of an examination board sitting, a record of the final result is entered, including details of Chair's action and referrals.

The system can be used to generate assessment information about a module cohort. It is planned to adapt the system to provide modular tracking of student assessment as the University's degree programmes become more clearly defined in modular terms within the SIS. Each module mark would be entered on to the system and automatic aggregation would take place to produce a module and final mark. It has been identified that the system would operate more effectively if all the Faculties had a standard spreadsheet which could then be networked and automatically accessed and updated.

The SIS is used as an index to a file store of complete information records on modular provision, which is held outside of the formal SIS system. This is used to retrieve information for use in the production of course flyers and handbooks. It is intended to combine these two systems in the future.

Case Study 12 - University of Edinburgh, School of Biology (27)

In 1988 the unit set up to co-ordinate the biology teaching programme, was re-organised as the Biology Teaching Organisation (BTO). There were four years of undergraduates doing a variety of course towards twelve different Honours degrees. The main objective of the unit was to develop the curriculum and to organise the administration of courses. Between 1988 and 1992 the number of undergraduates entering biology doubled from 250 to 500. Previously the administration for these students would have been spread across twelve different departments, seven in Science, four in Medicine and one in Social Science.

The seven Science departments were amalgamated to create the Division of Biological Sciences. The four medical departments remained in the Faculty of Medicine, but their teaching was organised through the BTO. The University's central record-keeping system proved to be inadequate for the modular Biological Sciences degree programme. Student records had a low priority, and due to technical problems the BTO was unable to tap into the central database which made administration slow and difficult. Because the BTO found that information was needed faster than the central system could provide, it created its own fast computerised system.

The system developed is a database linked to an optical mark reader (OMR) and a bar-code reader, and it is such that it can be altered and customised to suit an individual department's requirements. The database was developed to cope with both the increased number of students and the increasing movement of students between the now modularised courses. It can be interrogated in a variety of ways depending upon the type of information required. The BTO decided to use an OMR to enter information on to the database as this removes the need to key in the data, which is where most errors arise, and because it is very fast. One member of staff developed the idea of using bar-coding to identify students and course codes for entering this information into the database, items which are least suited to input via the OMR. A bar-code encrypts each student's seven-digit enrolment number and can be printed on to labels or directly on to OMR MCQ answer sheets. This method has been used effectively to register new students, enrol them on courses, and in marking MCQ examinations, and the combination of bar-coding and optical mark reading has proved very successful.

Bar-coding and OMR is used to produce course lists. Each student marks on an OMR sheet, which has their personal bar-code, those courses which they intend to take, and this data is scanned into the computer (at a rate of >500/hr), which can then sort the students according to course and print-out course lists. As students change courses the database can be easily updated.

The database used is *Borland's Paradox*. All the data are held on a fileserver and there is a separate directory for each year. Student records are held containing names, enrolment numbers, and modules being taken, and there is also a timetabling file which holds each student's name, number, courses taken, and details of tutorial, seminar and practical sessions. An update column exists to show when alterations were last made.

The system allows the extraction of individual student records. It can compile lists of which students do which course and what their timetabling restrictions are. This is used to help tutors rapidly assign the new intake of students to practicals and tutorials, and it is also used to spot students doing odd combinations of subjects and to prevent over-subscription of courses. It is invaluable at the start of each session and, when used with pre-registration for the next session, gives advance warning to course organisers or the likely number of students on their courses.

The A4 sheets for OMR scanning are printed (black-on-white) in-house, which reduces the cost. Different styles of sheets are used to register new arrivals, to indicate choice of course, to confirm choice of course for returning students and mark answers to MCQs and course feedback questionnaires. There is currently a move towards anonymous marking which will be aided by the existing bar-codes for each student number. The bar-coding is read by a simple and cheap hand-held wand rather than some of the more sophisticated readers because this enables the reading of small bar-codes. If there is a problem with the system, or mistake on the bar-code, it is possible to enter the student information manually through the keyboard.

Chapter Eight

Proceed - But with Caution!

Introduction

In this section are some words of warning about the use of technology in the assessment of students. The design of an assessment task is something, which involves careful consideration. Assessment should be integrated with the course aims and objectives and it is therefore important to know what each task is assessing. Much assessment in higher education has been repetitive. For example, by giving students only essays as a form of assessment, the same set of skills is being assessed repeatedly. It is a more thorough and relevant sampling of a syllabus to use a variety of assessment methods, with different methods developed to assess different types of learning. Appendix 2 illustrates the taxonomy of examination techniques extracted with permission from the Engineering Professors' Conference (1992). This is a useful and illuminating guide to the wide range of different assessment methods and their potential to test different types of learning.

With the increasing use of CAL and CBL in higher education there has recently been a move towards automated assessment. There is a tendency to jump on the bandwagon of technology with an enthusiasm, which leads to high expectations, which often cannot be met. It is crucial to realise that CAL and CBL applications do not have to be assessed using computer-aided assessment. This applies to other technologies such as multimedia, hypertext and so on. These applications are presently assessed in three ways:

- assessment is integrated in the CAL/CBL package;
- a separate software package is used to assess what the student has learnt from the CAL/CBL package;
- a paper and pencil test is used.

There is a danger in assuming that technology-based assessment will solve all problems. Assessment needs to be appropriate for the objectives of the course, and the choice of assessment method should follow the identification of learning outcomes and criteria. Before embarking on a programme of technology-based assessment the purpose must be clearly identified and the appropriateness of the technology it is intended to use.

There are several considerations, which apply to the implementation and use of technology-based assessment systems. The main considerations before designing and implementing such a system are:

- **Time** The amount of time taken to design and create technology-based assessment is often under-estimated. Programming software can be very time-consuming, even to produce a short programme. The time taken to create the assessment must be offset against the timesaved in marking and administration. Using existing packages, such as *Examine*, *Question Mark* cuts down design time.
- **Resources** Lack of resources other than time is often a great hindrance to the development of technology-based assessment. It is important to know what is available: hardware; software; technical expertise; money; space and equipment, before embarking on designing an ambitious assessment system that may be impractical.
- **Shelf-life** Will what is created soon be obsolete? This could happen because the discipline has changed its emphasis, or advances have been made in the area of technology used.
- **Portability** The system should be adaptable to take account of changes in subject content and new objectives. The chances of your system of assessment becoming obsolete may be reduced through a high level of portability.

If there is the time, resources and capacity to design a portable technology-based assessment system, there are number of initial questions which need to be answered before you begin. Some of these questions relate specifically to technology-based assessment, others are more general.

- **What is to be assessed?**

  Does what you are assessing meet the aims and objectives of the module/course? Have you specified criteria, assessment outcomes?

- **What type of assessment is it?**

  Is it formative or summative or both? Does it contribute to a final assessment, or not?

- **Should the assessment system enable tailoring of assessment tasks?**

  Tailoring an assessment task means that as an individual completes a section of the assessment task, it is marked automatically and the next section presented is partially dependent on previous performance. This saves time, as students can avoid undertaking unnecessary assessment. Pathways through the assessment system can be monitored to provide useful feedback for students and tutors.

- **Does it replace or supplement another non-technological assessment methods?**
  Using technology-based assessment to supplement or replace existing assessment with increased efficiency is useful. Using it to duplicate another assessment is not.

- **Is it efficient?**
  Does it give you savings in, for example: time (less marking); money (spent on paper, ink, administrative costs); accuracy (fewer clerical errors, less checking needed); faster and better feedback?

- **Is it effective?**
  Is it really the most appropriate method of assessment - or is it just an excuse to have a computer programme, multimedia application, video camera in your department?

- **Does it fit in with the rest of the course?**
  How does the system integrate with existing lectures, seminars, assessments?

- **Does it support further learning?**
  If used for formative assessment, your system should encourage students to refer to texts, CAL, CBL materials to further their learning.

- **Is it within your students' capability?**
  Are your students sufficiently skilled to use the system? Will they have to be provided with training? Will it be necessary to create guidelines for navigation around the system?

- **Will your students like it?**
  There may be resistance from students when introduced to technology-based assessment methods. It is important to remember that the majority may have no experience of this type of assessment and may need persuasion and training.

- **Is there adequate support?**
  If someone else is writing the programmes, will that person still be available in a year's timewhen something goes wrong, as it often does? Will there need to be anyone present to help students when they are using the assessment system?

- **Does it simply shift the workload to someone else?**
  It may be that the assessment method means less marking for you but more administrative work for a colleague. That may be your objective, but it is important to be clear about the *total* savings to be made. The system must be considered as a whole.

- **What do colleagues think and has the system been sufficiently tested?**
  It will be useful to consult colleagues, who may have some helpful suggestions. It will also help to consult students, to ask their opinions and to test the system out before putting in into practice.

- **Is it really necessary?**
  Is it really a more efficient and effective assessment procedure than the one operating at the moment?

The questions outlined above need careful consideration before a technology-based assessment system is implemented. It is vitally important that any system is well integrated with existing lectures, seminars, practicals and with other assessment methods. Technology-based assessment should not be an 'afterthought', but should be appropriate and effectively integrated into course and assessment design. In Case Study 1, the use of technology enables fast, effective and meaningful feedback to be given to students through their tutorial system.

An integrated technology-based assessment system should encourage and motivate students to learn, not only from technology-based materials, but also from other resources. There is much potential, especially in computer-based materials to direct student learning towards appropriate learning resources. Case Study 9 gave an example of how students can be motivated to look to a variety of courses, through the use of CBL materials.

Assessment should also be flexible, giving students the opportunity not only to choose the pace at which they learn, but also the level at which they are assessed. Case Study 4 illustrated a learning environment in which this is possible. Students should, through the assessment process, be given the opportunity to manage and monitor their own learning. Technology-based assessment often provides opportunities for learning management and monitoring which would otherwise be difficult and time-consuming to achieve in a traditional system.

A technology-based assessment system should achieve all of the above to be effective, efficient and worthwhile.

Future developments

Technology-based assessment techniques are currently being researched and developed in a number of higher education institutions in the UK. Many of the projects under the Teaching and Learning Technology Programme (TLTP) are creating courseware, which will include technology-based assessment techniques. Some areas under development in higher education include the use of modified essay questions, artificial intelligence, virtual reality, interactive video and computer simulation. The development of national and international networks, such as JANET and World Wide Web present rich opportunities for the development of teaching, learning and assessment in higher education.

A computer-based assessment project has recently been set up at Glasgow Caledonian University. It is hoped to produce a campus-wide system with two interfaces to allow tutor and student access. The system will hold a bank of questions, which will be used by students in Engineering as practice for examinations. Questions will be diagnostic of knowledge and understanding, and the system will carry out diagnostic testing, helping students to realise their strengths and weaknesses.

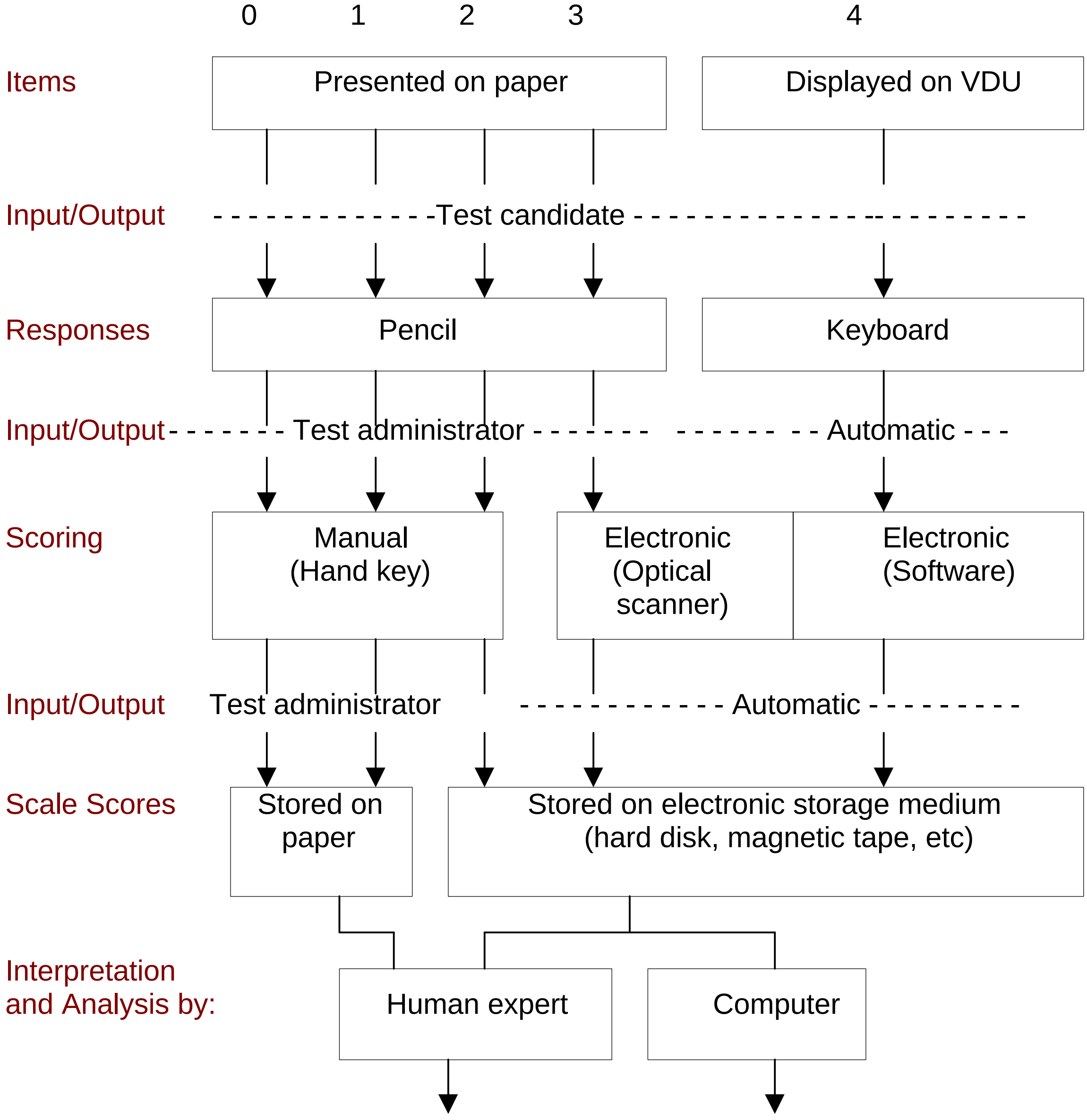
Perhaps the biggest hindrance to the development of technology-based assessment systems is the fact that academics understandably are most confident when they are using techniques and tools which they themselves have developed. (This is sometimes referred to as the 'not invented here syndrome'.) Whilst unwillingness to adapt and adopt others' systems may appear to be unreasonable, there is an advantage in developing your own system, in that you know it inside out, and have the ability to alter component parts. It will also fit the purpose for which it was designed. However it is impractical and wasteful to expect everyone to develop their own system and it is hoped that the future will bring the development of technological assessment shells which are portable and may be customised to meet individual requirements. There is also some merit in developing systems which all present the same structure and follow similar procedures - it is after all the students' performance on their course which is being assessed and not their ability to navigate through a variety of different assessment packages. Packages such as *Authorware Professional*, *Toolbook* and so on bring the ability to design and create technology-based assessment tasks closer to the hands of individual academics

Please write to us about any projects or technology-based assessment methods and procedures, which you have developed, or are developing, so that we can include them in any future editions of this survey. *Joanna Bull, Project ALTER, USDU, Level Six, University House, Sheffield S10 2TN. Tel: 0742 750820, Fax: 0742 728705, email: j.bull@sheffield.ac.uk*

Appendix 1

Various degrees of automation involved in the different modes of test administration and scoring

Test Administration Mode



Mode 4 Fully automated test administration and scoring. The items are presented by the computer; responses are entered directly into the computer, which calculates scale scores and maintains a database sufficient to provide access to other software dealing with the analysis and

interpretation of results.

Mode 3 This is similar to Mode 4, but the test items are presented in their 'traditional' paper format, while the responses are entered directly into the computer.

Mode 2 Here the complete test is administered 'off-line' using paper and pencil, but the response sheet is scored electronically (using an optical scanner) with the data being transferred directly to a computer. These data then have access to the software needed to compute scale scores, carry out profile analyses, perform computer-based test interpretation and so on.

Mode 1 As for Mode 2, but with the scoring done by hand, using scoring key. The scale scores and clients' details are hand entered on to a computer and can access the same analysis and interpretation packages (see below) as Modes 2 to 4.

Mode 0 The traditional 'non-automated' form of paper-and-pencil administration with hand scoring and all analysis and interpretation done 'by hand'.

[Extracted with permission from Bartram, D. (1989) 'Computer-based Assessment' in Herriot, P. (ed) Handbook of Assessment in Organisations. *Chichester: John Wiley & Sons Ltd.]*

Appendix 2 (a)

A Taxonomy of Examination Styles

This taxonomy compares examination styles according to their potential to test levels of learning of different kinds. It offers a basis for selection of assessment methods to match specified educational aims. It makes a number of assumptions, based largely on experience, which seem to be generally valid. However, it is possible to change the value given for 'potential for assessment' of some items in the taxonomy by the careful design of questions.

Knowledge means information which has been memorised and can be recalled.

Skills I means 'measurable' skills which can be judged objectively - like Maths calculations, spelling and grammar, computer skills, etc.

Skills II means 'complex' skills which require judgement by examiners - like communication skills, interpersonal skills, leadership skills, etc.

Reliability refers to the likelihood of different examiners giving the same mark.

Robustness refers to the resistance of the method to abuse by students (eg by cheating, plagiarism, unauthorised group activity, etc).

References and Contact List

Chapter two

1 **Commercial packages**
NCS Services (UK Ltd.)
Hellaby Business Park
Hellaby
Rotherham
S66 8HN
Tel: 01709 700400

2 **Data & Research Services Plc**
Sunrise Parkway
Linford Wood
Milton Keynes
MK14 6LR
Tel: 01908 666088

3 **Question Mark**
Mr M. Silver
Question Mark Computing Ltd
5th. Floor
Hill House
Highgate Hill
London N19 5NA
Tel: 0171 2637575

4 **Hypertext self-assessment questions**
Email: tltsn@ukc.ac.uk
Email: tltsnweb@icbl.hw.ac.uk

Examine Multiple-Choice and Short Answer Quiz Authoring System
Bio-Informatics Research Group
Department of Life Science
University of Nottingham
University Park
Nottingham
NG7 2RD
Tel: 01159 515151 ext. 8508/8507

6 **Question Handler**
Institute of Computer-based Learning
Queen's University of Belfast
Belfast
Northern Ireland
BT7 1NN
Tel: 01232 245133

7 **Case Study 1 - MCQs,**
Department of Biochemistry, University of Leeds
Dr Andrew Booth
BioNet Teaching and Learning Technology Programme
Department of Biochemistry and Molecular Biology
University of Leeds
Leeds
LS2 9JT
Tel: 01132 431751
Email: a.g.booth@leeds.ac.uk

Chapter Three

8 **Case Study 2 - THESYS**
Professor Richard Gentle
Department of Mechanical Engineering
The Nottingham Trent University
Burton Street
Nottingham
NG1 4BU
Tel: 01159 418418

Chapter Four

10 **Dr J. Pedler**
OSCAL
Leeds Dental Institute
Department of Dental Surgery
Clarendon Way
Leeds
LS2 9LU
Tel: 01132 440111

11 **Problem-solving - 'Energy Targeter'**
Mr Ian Frame
School of Built Environment
Anglia Polytechnic University
Victoria Road South
Chelmsford
Essex
CM1 1LL
Tel: 01245 493131 ext 3291
Email: c.i.frame@anglia.ac.uk

12 Case Study 3 - Computer-assessed practical work

Dr Peter Grebenik
Department of Biology and Molecular Sciences
Oxford Brookes University
Gipsey Lane
Headington
Oxford
OX3 0BP
Tel: 0865 741111
Email: pgrebenik@brookes.ac.uk

13 Case Study 4 - CALM Project

Professor C. E. Beevers
Mathematics Department
Heriot-Watt University
Riccarton
Edinburgh
EH14 4AS
Tel: 0131 449 5111
Email: c.c.beevers@hw.ac.uk

14 CALMAT

Dr Jean Cook
CALMAT
Department of Mathematics
Glasgow Caledonian University
Cowcaddens Road
Glasgow
G4 0BA
Tel: 0141 3313000 x3053
Email: j.cook@gcal.ac.uk

15 Submission and testing of computer programming coursework

SUBMIT

Dr John Davy
School of Computer Studies
University of Leeds
Leeds
LS2 9JT
Tel: 01132 431751

16 Computer checking of programs and their design

P L Fazackerley, P Halstead,
Computing Department
The Nottingham Trent University
Burton Street
Nottingham
NG1 4BU
Tel: 01159 418418

17 Case Study 5 - CEILIDH
Steve Benford, Edmund Burke, Eric Foxley
Department of Computer Science
University of Nottingham
University Park
Nottingham
NG7 2RD
Tel: 01159 515151

18 Case Study 6 - Computer-assessed examinations
Dr Pamela Williams
Department of Italian
University of Hull
Hull
HU6 7RX
Tel: 01482 346311

Chapter Five

19 Case Study 7 - video case study
Leslie Chadwick
Financial Management and Accounting
Management Centre
University of Bradford
Richmond Road, Bradford, BD7 1DP
Tel: 01274 232323

20 Use of computerised statistics packages in traditional examinations
Tony Downing/Dr K. Knott
Department of Psychology
University of Newcastle upon Tyne
6 Kensington Terrace
Newcastle upon Tyne
NE1 7RU
Tel: 0191 222 6184/7519

21 Case Study 8 - Computerised examinations
Peter Keenan
Department of Management Information Systems
University College Dublin
Dublin 4
Ilrelan
Tel: 00353 1706 7777

22 Guide hypertext project
Dr Pamela Williams
Department of Italian
University of Hull
Hull
HU6 7RX
Tel: 01482 465993
Email: p.a.williams@selc.ac.uk

23 Case Study 9 - MacCycle

Dr Jim Aiton
School of Biological and Medical Sciences
Bute Medical Buildings
University of St Andrews
Fife
KY16 9TS
Tel: 01334 476161
Email: jfa@st-and.ac.uk

24 Computerised module management system

Anne Pollock
Quality Assurance Unit
Napier University
Craiglockhart Campus
219 Colinton Road
Edinburgh
EH14 1DJ
Tel: 0131 444 2266

25 Case Study 10 - The Open University

Mr Ben Palmer
Deputy Academic Registrar
(Course Presentation and Examination)
The Open University
Walton Hall
Milton Keynes
MK7 6AA
Tel: 01908 274 066

26 Case Study 11 - Leeds Metropolitan

University, Student Information System

Jim Sinclair
Room C212
Leeds Metropolitan University
Calverly Street
Leeds
LS1 3HE
Tel: 0532 832600 ext. 3103
Fax: 0532 833115

27 Case Study 12 - University of Edinburgh, School of Biology

Dr Jeff Haywood
Department of Biochemistry
University of Edinburgh Medical School
Hugh Robson Building
George Square
Edinburgh
EH8 9XD
Tel: 0131 650 4131
Email: J.Haywood@ed.ac.uk

Bibliography

Bartram, D. (1989) Computer-based Assessment, in Herriot, P. (ed) Handbook of Assessment in Organisations. Chichester: Wiley.

Bell, R. C. (1980) *Problems in Improving the Reliability of Essay Marks.* Assessment and Evaluation in Higher Education 5, 254-63.

Bliss, J. and Ogborn, J. (eds) (1977) *Students Reactions to Undergraduate Science.* London: Heineman.

Brewer, I. M. (1985) *Learning More and Teaching Less.* London: SRHE.

Brown, G. A. and Pendlebury, M. (1992) *Assessing Active Learning.* Sheffield: CVCP/USDU. (See inside front cover for details.)

Boud, D. (1987) Implementing Student Self-assessment. Sydney: HERDSA.

Crooks, T. (1988) *Assessing Student Performance.* Sydney: HERDSA.

Daines, J. M. (1986) *Self-assessment in a Laboratory Course on Dispensing.* Unpublished PhD, University of Nottingham.

Educational Development Team (1993) *Innovations in Teaching 1992-3,* The University of Hull.

Entwistle, N. J. (1992) The Impact of Teaching on Learning Outcomes in Higher Education: A Literature Review. Sheffield: CVCP/USDU.

Falchikov, N., and Boud, D. (1989) *Student Self-Assessment in Higher Education: a Meta-Analysis.* Review of Educational Research 59, 395-430.

Gibbs, G., Habeshaw, S. and Habeshaw T. (1988) *53 Interesting Ways to Assess your*

Students. Bristol: TES.

Harvey, L., Burrows, A. and Green, D. (1992) . *'Criteria of Quality.'* Birmingham: QHE.

Hegarty-Hazel, E. (1986) *Research on Laboratory Work*, in Boud, D., Dunn J. and Hegarty-Hazel, E. (eds) *Teaching in Laboratories*. London: SRHE/NFER-Nelson.

Knight, P. (1993) *'Measure for Measure: An Assessment Toolkit'*, Lancaster University, Unit for Innovation in Higher Education.

Olimpo, G. (1986) *The Computer in the Evaluation of Learning*. Genova: Istituto per le Technologie Didattiche.

Rowntree, D. (ed) (1987) *Assessing Students - How Shall we Know them?* London: Harper and Row.

Wilson, G. et al (1969) Examinations of Clinical Examiners. Lancet 92, 37-41.

Further Reading

Assessment

Boud, D. (1991) *Assessment in Problem-based Learning*. Assessment and Evaluation in Higher Education 13, (2) 87-91.

Brown, S. and Dove, P. (eds) *Self and Peer Assessment*. Birmingham: SCED Paper 63.

Matthews, J. (1981) *The Use of Objective Tests*. Lancaster: University of Lancaster Press Revised edition.

Employment Department (1993) Assessment Issues in Higher Education.

Technology

Anandam, K. (1989) *Transforming Teaching with Technology: Perspectives from Two-year Colleges*. EDUCOM: Strategies Series on Information Technology. McKinney, Texas: Academic Computing Publications Inc.

Bartram, D. and Bayliss, R. (1984) *Automated Testing: Past, Present and Future*. Journal of Occupational Psychology, 57, 221-37.

Benford, S., Burke, E. and Foxley, E. (1992) *Automatic Assessment of Computer Programs in the Ceilidh System*. Submitted to IEEE-CS International Software Metrics Symposium.

Brook, A. and Ponicter, N. (1990) *Practical Technology Useful in Authoring Technology-based Learning Material*. Journal of Computer-based Instruction, 17 (2) 53-60.

Butcher, P. G. and Greenberg, J. M. (1991) *Educational Computing at the Open University: the Second Decade*. CALRG Report, No. 86, Open University, Milton Keynes.

CTI Centre for Mathematics and Statistics (February 1993) *Quarterly Newsletter*. Birmingham University.

Dann, P. L., Irvine, S. M. and Collis, J. M., (1991) *Advances in Computer-based Human Assessment*. Kluwer.

Gotz, K. (1991) *Interactive Learning with the Computer*. Education and Training Technology International 28, (1) 7-14.

Graves, W. H. (1989) *Computing Across the Curriculum: Academic Perspectives*. EDUCOM: Strategies Series on Information Technology. Mckinney, Texas: Academic Computing Publication Inc.

International Yearbook of Education and Instructional Technology. London and New York. Kogan Page/Nichols: (Published annually for the Association for Education and Training Technology.)

Sportier, J. W. (1989) *Facilitating Academic Software Development*. EDUCOM: Strategies Series on Information Technology. Mckinney, Texas: Academic Computing Publications Inc.

Tuckey, C. J. (1992) *Uses of New Technology in Higher Education - Guiding Principles:* ICBL Reports, No 1. Heriot-Watt, Edinburgh: Institute for Computer-based Learning.

Verth, P. B. Van, (1985) *A System for Automatically Grading Program Quality*. SUNY (Buffalo) Technical Report.

Zin, A. M. and Foxely, E. (1991) *Automatic Program Quality Assessment System*. Vidyangar, India. (In proceedings of the IFIP Conference on Software Quality).