

---

# Revisiting objective tests: a case study in integration at honours level

Pauline Maclaran\* and Alan Sangster\*\*

\*Department of Marketing, De Montfort University

\*\*The Open University Business School

email: a.sangster@open.ac.uk

---

*This paper examines the background to computer-assisted assessment and shows how certain misconceptions or 'myths' have arisen around its use. It then discusses an actual implementation of computerized multiple-choice question (MCQ) tests, addressing both the main theoretical issues, and the practicalities of the design and administration process. It confirms that honours-level learning can be appropriately assessed using summative computer-based objective tests, not just in the eyes of the adopting academic, but also in the eyes of the students. Care should, however, be taken to adopt a flexible implementation that is responsive to unforeseen problems.*

---

## Introduction

Multiple-choice questions (MCQ) are widely accepted in the United States as a mode of assessment in undergraduate courses. In an environment of ever increasing class sizes, they provide a faster way to assess large groups of students, whilst also providing a way to measure deep understanding. However, in the UK there is scepticism from lecturers about the use of MCQ tests and, perhaps more significantly, from students themselves. This frequently means that someone using MCQs as a means of assessment will find both a lack of support from colleagues and resistance from students.

When MCQ tests are computerized, the introduction of this additional non-traditional element in the assessment process can compound this situation. This paper seeks to provide evidence to support anyone considering using this mode of assessment. A description is given of a case study of the use of computerized multiple-choice summative assessment for a class of 260 honours-level undergraduate students taking a marketing module.

---

## **An overview of computer-assisted assessment**

Much computer-assisted assessment (CAA) relies on the incorporation of MCQ tests. The preferred term for MCQs is 'objective tests', although they are, of course, only as objective as the person who designs them (Brown, Bull, and Pendlebury, 1997). Objective testing is a broader and more appropriate term given that much MCQ software incorporates more than just selection of one answer from four or five choices. This is certainly the case with the *Question Mark Designer* program used for this present case study.

One of the most common misconceptions that has occurred in relation to this type of testing is that it is less about student learning and more about lecturer convenience. It is frequently considered to be, at best, a practical way to manage larger classes and, at worst, an easy option for lazy colleagues. This 'easy option' myth has arisen mainly because of the emphasis throughout the 1980s and early 1990s on the use of computer-assisted learning (CAL) to manage larger classes. This criticism certainly has some foundation, for objective testing does provide a more efficient way to respond to larger class sizes. However, when effectively and appropriately integrated into the student learning experience, it not only has sound pedagogical validity, it can take as much, if not more, effort to administer than traditional approaches to assessing large classes. Appropriately applied, objective testing is certainly not the 'quick-fix solution' that some suggest (Bull, 1994).

More recently, there have been an increasing number of studies that investigate pedagogical issues such as:

- learning styles in relation to computer-based learning (Groat and Musson, 1995; Sangster, 1995; Valley, 1997);
- methodologies for integrating CAL into the curriculum (Persico, 1997; Conole and Oliver, 1998);
- the relationship of theories of learning to courseware design (Somekh, 1996).

These studies emphasize the importance of a holistic approach to computer-assisted learning that considers overall educational goals, and that is designed as part of a comprehensive learning plan and not just a stand-alone or ad hoc add-on.

A further benefit of learning technologies, such as CAA, is that they offer alternative avenues of tuition to students (Keady, 1997). This flexibility is particularly useful for students at a distance, or those who require additional support. It also allows and encourages independent learning. With so much emphasis on a progression from high teacher support levels to increased student empowerment (Whalley, 1995) CAL can enable students to learn material at their own pace and in their own preferred ways.

## **Integrating MCQs into the curriculum**

Students tend to focus on those aspects of a course that they perceive will contribute to their overall marks (Brown et al, 1997). One major benefit of objective tests arises from their ability to test breadth of knowledge. Course developers who take advantage of this feature will often find their students read more widely around the subject than they might do in a conventionally delivered module in the same subject. The introduction of computer-based objective tests as a primary mode of course assessment on the B.Sc.

Management marketing module was devised to achieve three goals:

- to encourage the development of a breadth of subject knowledge in the students;
- to encourage the students to take responsibility for their own learning;
- to encourage them to utilize a Website entitled the *Marketing Learning Centre* (Catterall and Ibotsen, 1998) accompanying their core textbook (Dibb, Simpinkin, Pride and Ferrell, 1997).

The Website (Figure 1) offers the following learning support: a bank of MCQs to accompany each chapter, Internet exercises to learn about marketing organizations and case companies, links with Canadian and US Marketing Learning Centres and a glossary of terms.

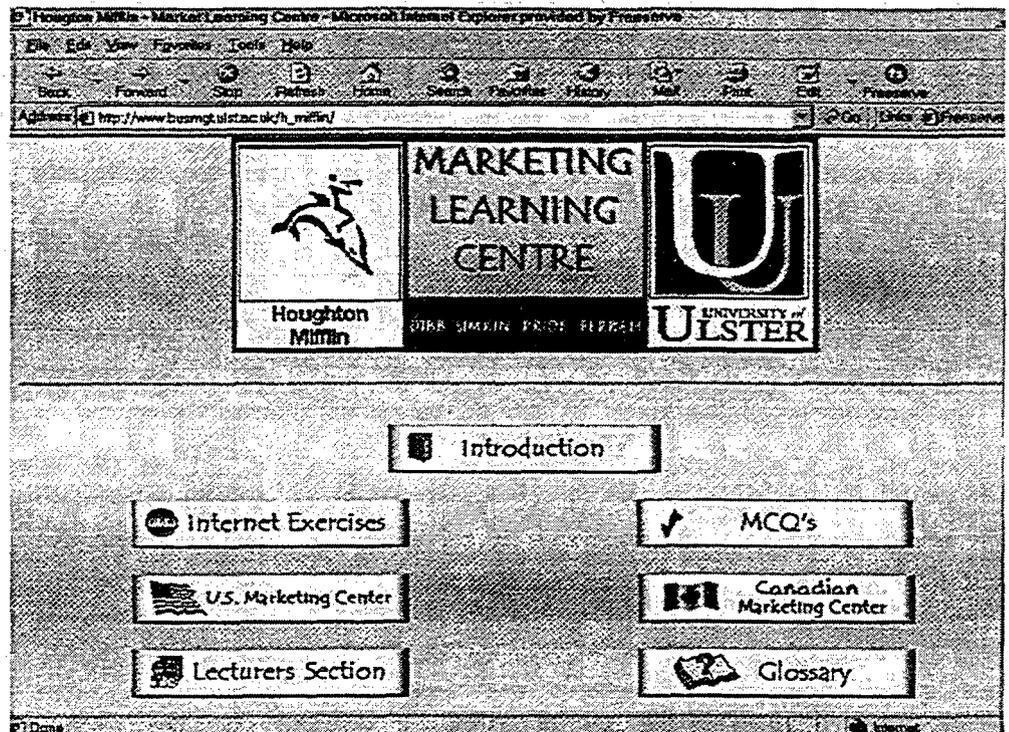


Figure 1: The Marketing Learning Centre Webpage.

One resource of the *Marketing Learning Centre* that the students particularly appreciated was the links to the Web pages of companies such as Kelloggs, Lego, Disney, and McDonald's. In accessing these sites students were able to study how the companies were incorporating their Websites into their overall marketing strategy. Moreover, these exercises provided a starting point for students to begin to explore the Web and discover other links for themselves.

Students were allocated a supervised one-hour computer laboratory session each week, working in groups of around sixty students. In these sessions, they accessed the *Marketing Learning Centre* and worked through the learning support materials in whatever manner

they chose. They could work with each other or alone, in sequential order if they wished, or they could alternate between topics and relevant Websites, etc. The students were also expected to spend an additional four to six hours per week on the material outside the formal class sessions.

This flexibility in the way the students were encouraged to approach their learning was important, particularly to acknowledge the social aspects of learning as put forward by Bruner (1966) and Vygotsky (1986). Those that needed support from their tutor also benefited from the greater individual access to the tutor resulting from the increased freedom this undirected approach to student learning in the computer labs gave the tutor to have one-to-one contact as individual problems arose. A third benefit arising from the design of this module was the confirmatory support students could obtain from the bank of MCQ tests in the *Marketing Learning Centre* (see Figure 2). This formative assessment provided quick feedback on their progress, feedback that would have otherwise not been so promptly available, particularly given the large class size.

Below is a series of 20 Multiple Choice Questions which you must select an appropriate answer for each. To answer a question simply use your left mouse button to click the small radio button that appears at the start of each individual question, you should then see your selected answer has been highlighted by a small black dot to change your answer simply click on your new answer and this should deselect your old answer and select your new answer.

Once you are certain that you have answered ALL the questions correctly, click on the button after the last question that states MARK ANSWERS this will lead you to another page which shows how successful you have been - Good Luck!!

Please ensure that you fill in the appropriate details below:

Student Details:

Name:

Institution:

1) Marketing is best defined as:

A)  matching a product with its market.

B)  promoting and selling products.

C)  facilitating exchange relationships.

D)  distributing products at the right price to stores.

2) The expansion of the definition of marketing to include nonbusiness activities adds which one of these examples to the field of marketing?

A)  Proctor and Gamble selling toothpaste.

Figure 2: An example of the Marketing Learning Centre MCQ tests.

A crucial factor in motivating the students was that their learning would be tested by two summative computer-assisted objective tests, in the sixth and tenth weeks of the twelve-week course. These were specifically designed to reflect the majority of topics covered during the two hours a week of lectures and one hour a week of computer sessions. Thus, the design of the course was such that the two objective tests played a pivotal role in driving the learning of the students.

Furthermore, because these tests were assessing specific knowledge about the principles of marketing and their applications, the final written examination was designed to take a more contextual and holistic approach than would otherwise have been appropriate. The examination comprised two compulsory questions. The first was to put together a mini-case study to illustrate a marketing-orientated company. The second was to discuss how the Internet could be used in a firm's marketing strategy. Students were advised early in the course of the nature of these questions, ensuring they had an incentive both to pay attention to the various case studies they encountered and to make use of other material that they accessed through the Web, either directly or via the *Marketing Learning Centre*. It can be seen, therefore, that both the formative MCQs and the summative objective tests were very much designed as an integrative formative part of the overall student learning experience in this module.

### **Design of MCQs using Question Mark Designer**

The software program used for the design of the objective tests was *Question Mark Designer*. It allows for a flexible approach to question design and offers a variety of question types and formats, including the ability to introduce graphic images. The main question types include explanation, multiple choice, text match, multiple response, numeric and selection. There are different formats within each of these categories. For example, within 'selection' there are options for true/false/don't know, agree/disagree, ranking, and matching statements. This variety of forms is important if different levels of learning are to be integrated through this approach into the assessment process.

In terms of Bloom's taxonomy of learning (Bloom, Englehart, Frost, Hill and Krathwoh, 1956) – knowledge, comprehension, application, analysis, synthesis and evaluation – teachers often ask questions only in the 'knowledge' category. Kehoe (1997), for example, estimates that this occurs 80–90 per cent of the time and suggests that a higher-order level of questions should be used, particularly in the writing of MCQs.

Not surprisingly, as the main purpose of the *Marketing Learning Centre* is to support a build-up of knowledge around topics in the textbook, the question banks it contains fall mainly into the 'knowledge' category, i.e. they rely on remembering, memorizing, and recognizing. As a result, while they are very useful in supporting and enhancing student learning, their scope is too limited for them to be suitable for use as summative assessment on a module of this type. For summative assessment, it was seen as important in the context of this honours-level module to avoid a predominance of question types and formats that require only 'knowledge' responses. This meant moving away from conventional multiple-choice and text match designs towards the introduction of question types that encourage deeper levels of analysis and evaluation such as 'multiple response' or 'selection' type questions. The latter type has one particularly useful format called 'Web' that allows the presentation of information, for example a mini-case study or detailed description of a company or event. Questions can then test how this information has been comprehended, analysed, synthesized and evaluated, thereby extending question categories as described in Bloom taxonomy.

This also obviates one of the principal criticisms levelled at MCQs, that they will encourage surface learning in place of a deeper approach (Entwistle, Thompson and Tait, 1992). This is

a contentious point. Webb (1997) goes so far as to suggest that the deep-surface dichotomy is little more than a ship of convenience and argues that it is a question of personal perspective and culture as to whether 'deep' is better than 'surface', or vice versa. Whatever the merits of these arguments, there is no reason why MCQs cannot be designed to test at the higher levels of Bloom's taxonomy and, therefore, encourage learning beyond the level generally considered 'surface' (see, for example, Fogarty and Goldwater, 1992).

Apart from scope in question design, *Question Mark Designer* provides good facilities for question evaluation. Once a test has been completed, the program will provide collated scores and a comprehensive question and answer analysis. Each question can be analysed in terms of its facility and ability to discriminate between students. For example, the analysis places questions on a continuum running from zero to one. Any question lying too close to either end of the scale is either too difficult (close to zero) or too easy (close to one). Questions can then be revised in the light of this analysis.

There are other useful program features, such as randomization of questions, several feedback alternatives, restrictions on test time, password control, and options for modifications in scoring, including negative marking.

### **Main implementation difficulties**

The first test in the sixth week comprised thirty questions to be answered in twenty minutes. The tests were held in the students' normal computer sessions and took place over a one-week period. Overall, there were four test sessions with around sixty-five students in each. The difficulties that were experienced during the actual implementation of the first test can be broadly divided into two categories – preparing the students for the tests and the process management of the taking of the tests.

#### **Student preparation**

Initially students were unconvinced that objective tests offered a valid way to assess their knowledge. Some time had to be spent ensuring that students would 'buy-in' to this type of assessment. This was done in several ways:

- Some of the early resistance disappeared as students gained experience of testing their knowledge through the *Marketing Learning Centre*. They liked being able to obtain immediate feedback on their progress. Where they worked in a group, a competitive aspect – albeit light-hearted – was evident.
- The integrative aspects of the course were explained in detail to the students, and frequently repeated. The fact that studying for these summative MCQ assessments would also help them prepare for the end-of-module written examination meant that the examination would be less stressful than usual and this was seen as a bonus by many of the students. In addition, the knowledge that they would ultimately have the chance to express themselves in a more traditional, essay-style format offered them further reassurance that they were not being disadvantaged by this non-traditional approach to their assessment.
- To ensure that the students were prepared for the more complex question types used in the summative objective tests, they were given demo disks with sample *Question Mark Designer* tests to practice upon.

- The introduction of negative marking for certain questions helped allay fears that lazy students could just be lucky on the day. This was introduced in respect of questions that required true or false answers and therefore offered a 50 per cent chance of a guessed answer being correct. A 'don't know' category was included that scored 0 marks. In the event that true or false were wrongly selected, the candidate was awarded minus 1.
- Negative marking is a controversial aspect of objective design (Turton, 1996). In this instance it was introduced on seven questions each with four statements to be evaluated. When the impact of this negative marking was analysed, it was found that no student had lost more than three marks as a result of its being used. Overall, negative marking reduced the mean mark for the first summative objective test by 1 per cent.

Despite all this preparation, the students were insufficiently prepared for the actual question formats of the first test. Many students appeared unable to move from the 'knowledge' type questions in the *Marketing Learning Centre* to the summative objective test's more demanding level of questions.

Time pressure was an additional factor that caused concern and for which they had no previous practice. While around a third of the class scored highly (at the upper-second/first class honours level) approximately a quarter failed the test altogether. Many of those in the middle range (lower-second/third class categories) felt they should have achieved higher scores.

The spreading panic and loss of confidence in the tests was only averted by introducing a third test and announcing that only their two best marks would count towards their final mark. This alleviated fears and meant that those who had done badly had a possibility of discounting their first mark altogether. With hindsight this notion of 'two out of three' would have been a good idea in any case and might have assisted the initial 'buying in' process.

In addition, the demo MCQs the students had been given to practise using the testing software were generic ones provided with the software. This did not appear to have had the desired effect of ensuring familiarity with the testing software, possibly because of the lack of relevance of the subject matter to their course. To address this, disks containing three short objective tests in marketing were created and provided to the students. Again with hindsight, these would also have been useful in the early stages to encourage an early flexibility in question responses and to stop the rut effect in student thinking caused by their extensive exposure to the knowledge-focused MCQs at the *Marketing Learning Centre*.

#### **Process management**

With such a large class, the process management of the first test was problematic and had several implications for the administration of the subsequent two tests. The four sessions with sixty-five students were spread through the week, from Tuesday to Friday. Four helpers were recruited to assist and supervise students. This was important to ensure that students did not copy the test to pass on to a friend. In the inevitable flurry of arrivals and departures, it still proved difficult to monitor students closely at the start and the end of the test.

In addition, students found it intimidating to be treated in the regimental that their numbers necessitated, particularly by other staff with whom they were not always familiar. Furthermore, when results from Friday's class (the last test) were compared with Tuesday's (the first), a 5 per cent increase across the range of marks was noted. Although the time pressure of the questions ensured that few were remembered, there was the worry that students who took the test early were being disadvantaged, even if only in terms of the time to prepare for the test.

For subsequent tests it was decided to assess all students on the same day in groups consisting of 25–30 students in a smaller computer laboratory. Ten sessions were run at forty-minute intervals throughout the day and this proved highly successful from the process management side. A disc was placed in each computer to record the scores of the ten or so students who used each PC. This saved time not only in copying the original test onto disks but also in collating the scores afterwards (a maximum of thirty disks versus sixty-five for the first test). It also meant that students would be given a computer number in advance of the test and go straight to their allocated seat upon their arrival in the computer lab.

Furthermore, the tests were permanently ready to run and this reduced the risk of disk copying that had been a major concern in the first test. The atmosphere in the lab was greatly improved from that in the first test. Students appeared more relaxed and they seemed reassured by the fact that their module lecturer conducted all the administration of the test. The same approach was also adopted for the third objective test.

### Student performance

The relative performance of the students in the three tests and in their exam can be seen in Table 1.

	Test 1	Test 2	Test 3	Overall	Exam	Overall
Number of students	255	246	179	257	251	251
60+	81	233	72	217	49	164
40–59	114	13	93	36	201	85
<40	60	0	14	4	1	2
Mean (per cent)	51.0	81.8	56.2	71.0	52.8	62.6

Table 1: Student performance.

There was clearly a problem with the level of mark attained in the second objective test. In order to restore students' confidence, it had been made appreciably easier than the first objective test and the negative marking dropped. Students commented that, as a result of their experiences of the first test, they had worked extremely hard for the second test. The combination of these two factors gave rise to the abnormally high marks. The third objective test was set to the same standard as the first test and the profile of the students across the three reported grades clearly shifted away from the fail band that had been so evident in the first test. Unfortunately, the second test had a significant impact on the overall score for the tests and this, in turn, resulted in a very skewed overall mark for the

course, despite the exam average being at approximately the same level as the averages for the first and second tests.

Obviously, finding the right level of test when objective tests are used in this way, is a non-trivial task. If it is too high, the students may lose confidence and may even withdraw from the course. If it is too low, marks across the entire course may be skewed and a final mark awarded that is not merited on performance across the course as a whole.

The marks achieved by the students on their other modules were similarly around 62 per cent on average, and the average mark achieved by students taking the module in the previous year was virtually the same at 61.9 per cent. Consequently, it was decided not to normalize or otherwise adjust the marks. In future, as both the level of test and the 'better' questions have been identified, this is not a problem that is likely to recur. However, it did give cause for concern among some colleagues who were already uncomfortable at an honours-level module being assessed in this way.

### **Conclusion**

Overall, students deemed the approach to the module a success, with some 68 per cent in their module evaluation forms mentioning the computer learning as one of the aspects they most liked about it. Similarly, they viewed the summative tests as being an appropriate form of assessment and one that had the added benefit of encouraging them to learn material as the module progressed instead of in a cramming exercise at the end of the module. However, the problems encountered illustrate the point made by Alavi, Yoo and Vogel (1997) that the integration of information technology into management education is not trivial; and that effective use requires a departure from traditional pedagogical models.

There are a number of lessons to be learned from this case study when preparing to adopt a similar approach:

- Prepare students to accept the concept of summative objective tests.
- Prepare students for computer-based summative objective tests by providing them with opportunities to gain knowledge and experience of using the software on domain-relevant questions before they have their first summative test.
- Integrate the tests into the overall design of the course.
- Ensure that different learning outcomes are tested by the examination.
- Experiment with question design and remember the learning levels you are testing.
- Plan the process management of the tests carefully and be prepared to alter it swiftly if problems arise.
- Above all, this case study confirms that computer-based objective tests can be used appropriately and effectively for summative testing at honours level.

---

## References

- Alavi, M., Yoo, Y. and Vogel, D. R. (1997), 'Using information technology to add value to management education', *Academy of Management Journal*, 40 (6), 1310-33.
- Bloom, B. S., Englehart, M. B., Frost, E. J., Hill, W. H. and Krathwohl, D. R. (1956), *Taxonomy of Educational Objectives: The Classification of Educational Goals*, Longman: New York.
- Brown, G., Bull, J. and Pendlebury, M. (1997), *Assessing Student Learning in Higher Education*, Routledge: London.
- Bruner, J. S. (1966), *Towards a Theory of Instruction*, The Belknap Press of Harvard University Press: Cambridge MA.
- Bull, J. (1994), 'Using IT for assessment: going forward', *Active Learning*, No. 1.
- Catterall, M. and Ibotson, P. (1998), *The Marketing Learning Centre*, [http://www.busmgt.ulst.ac.uk/lh\\_mifflin/](http://www.busmgt.ulst.ac.uk/lh_mifflin/).
- Conole, G. and Oliver, M. (1998), 'A pedagogical framework for embedding C&IT into the curriculum', *ALT-J*, 6 (2), 4-16.
- Dibb, S., Simpkim, L., Pride, W. M. and Ferrell, O. C. (1997), *Marketing*, Third European Edition, Houghton Mifflin: Boston.
- Entwhistle, N., Thompson, S. and Tait, H. (1992), *Guidelines for Promoting Effective Learning in Higher Education*, Centre for Research on Learning and Instruction: Edinburgh.
- Fogarty, T. J. and Goldwater, P. M. (1992) 'Instructor control in an automated environment: a reconsideration with empirical evidence', *Accounting Education*, 1 (4), 219-310.
- Groat, A. and Musson, T. (1995), 'Learning styles: individualizing computer-based learning environments', *ALT-J*, 6 (2), 53-62.
- Keady, H. (1997), 'Using computer based learning materials on the undergraduate programme', *Nottingham Business School Teaching and Learning Group Conference*, 5, 65-74.
- Kehoe, J. (1997), *How to Write Multiple Choice Questions*, <http://quizplease.com/lqhelp/lqphowto5.htm>.
- Persico, D. (1997), 'Methodological constants in courseware design', *British Journal of Educational Technology*, 28 (2), 111-24.
- Sangster, A. (1995), 'Objective tests, learning to learn, and learning styles', *Accounting Education*, 5 (2), 131-46.
- Somekh, B. (1996), 'Designing software to maximize learning', *ALT-J*, 4 (3), 4-16.
- Turton, B. C. H. (1996), 'A computer-aided continuous assessment system', *ALT-J*, 5 (2), 48-60.

Valley, K. (1997), 'Learning styles and courseware design', *ALT-J*, 5 (2) , 42–51.

Vygotsky, L. (1986), *Thought and Language*, MIT Press: Cambridge, MA.

Webb, G. (1997), 'Deconstructing deep and surface: towards a critique of phenomenology', *Higher Education*, 33, 195–212.

Whalley, B. (1995), 'Teaching and learning on the Internet', *Active Learning*, 2, 25–9.