2010

Experiences of Assessment Using Multiple Choice Questions On Advanced Modules Taken By Level 8 and Level 9 Engineering Students

Aidan O'Dwyer
Dublin Institute of Technology, aidan.odwyer@dit.ie

Experiences of assessment using multiple-choice questions on advanced modules taken by Level 8 and Level 9 engineering students

Aidan O'Dwyer,
School of Control Systems and Electrical Engineering,
Dublin Institute of Technology, Kevin St., Dublin 8
e-mail: aidan.odwyer@dit.ie

Abstract: This contribution evaluates the use of multiple-choice questions, in both formative and summative assessment modes, on advanced modules taken by Level 8 and Level 9 engineering students, over the past two academic years. Assessment data, and student experiences with the assessment methods, are reported and analysed.

1. Introduction

The use of multiple-choice questions in assessment has been reported extensively in the engineering education literature and in the wider educational context; for example, excellent on-line guides on the topic are available (e.g. Multiple Choice and Matching Tests, 2005). In the Irish context, disciplines that use multiple-choice tests include engineering (e.g. IIE Ireland diploma in industrial engineering programme www.iie.ie/education.asp), medicine (e.g. membership examination of the RCSI http://www.intercollegiatemrcs.org.uk/old/announcements_html), business (e.g. certified insurance practitioner programme http://www.insurance-institute.ie/downloads/education/2010/CIP.pdf) and language learning (e.g. Teastas Eorpach na Gaeilge http://www.teg.ie/english/info_advice_candidates.htm).

Multiple-choice questions can have two choices of answers (true/false), though, more commonly, four choices of answers are available. It is recognised that raw scores from these tests should not be used directly. The reason is that, for example, in a test where each question has four choices of answer, a student may know the answers for 20% of the questions and guess the answers correctly for one quarter of the rest of the questions, passing the examination. Scaling may be done using a probabilistic approach (Zhao, 2005, 2006) or a simpler approach (which employs negative marking). The scientifically sound probabilistic approach suggests that the optimum number of choice of answers for questions is 4. In addition, if the number of questions is greater than 18, for example, there is less than 1% probability of obtaining a scaled mark of 40% by pure guesswork. This probability falls to less than 0.01% if the number of questions set is greater than 48 (Zhao, 2005, 2006).

There is a vigorous debate in the literature about the role of multiple-choice questions in assessment. The advantages suggested for the use of such questions fall into three categories:

- They facilitate comprehensive student learning. The questions have the potential to cover the whole of the syllabus and they force the student, in principle, to learn all the taught material without exception (Excell, 2000). In a discipline-specific comment, Fenna (2004) suggests that multiple-choice tests are particularly desirable in engineering, to require the student to learn and correctly apply fundamental knowledge. In addition, multiple-choice tests are effective self-assessment tools (Davies, 1994; Azalov *et al*., 2004; McElroy *et al*., 2006). It is suggested that multiple-choice tests are especially suitable for knowledge-based

subjects which are well defined, do not change rapidly with time and have unambiguous right answers (Excell, 2000; Azalov *et al*., (2004); Zhao, 2005).

- Assessment efficiency. Well designed multiple-choice tests are an efficient means for the assessment of knowledge, analytical ability, language proficiency and numerical skills involving a large number of examinees (Zhao, 2005) and the tests are suitable where the relative competence of the examinees in a large sample size is to be assessed (Zhao, 2005). Automatic marking is possible (Excell, 2000) and results can be obtained quickly (Brown, 2001; Zhao, 2006), with test scores being reliable (Chang *et al*., 2007).
- They are amenable to analysis. Among the tools available are numerical measures of the quality of the individual multiple-choice question based on student selection of the answers and measurement of the ability of the question to discriminate between capable and weak students (Brown, 2001). A variety of other such metrics exist (DeSantis and McKean, 2003).

However, there are objections to the use of multiple-choice questions in assessment, which also tend to fall into three categories:

- General concerns about suitability. It is suggested, for example, that multiple-choice questions tend to address superficial facts, which may encourage learning of surface detail (O'Loughlin and Osterlind, 2007), though Struyven *et al*. (2006), among others, do not agree. O'Loughlin and Osterlind (2007) also suggest that even if the question is carefully worded, assessors cannot be sure that a student who answers correctly not only knows the correct answer but also understands the subject being examined and that students can select a correct answer for superficial reasons, such as selecting the answer through a process of elimination. However, it is possible to test student cognitive skills with properly constructed multiple-choice questions (Azer, 2003; DiBattista *et al*., 2004; Chang *et al*., 2007). It is also suggested that students find multiple-choice questions 'confronting' and would prefer to express themselves more fully (Brown, 2001). Perhaps the most extreme such comment is that of Azalov *et al*. (2004), who suggest that multiple-choice tests trivialise educational measurement.
- Concerns about suitability for assessing particular learning outcomes. For example, it is suggested that multiple-choice questions are not suitable for assessing numerical design exercises, a feature of many traditional engineering examination questions (Excell, 2000; Zhao, 2005).
- Structural issues. One concern is that students may select answers at random, though negative marking or other strategies can be used to reduce this (Excell, 2000). It is suggested that multiple-choice questions are 'very much more difficult' to write than descriptive questions (Brown, 2001). In addition, the number of possible questions rapidly becomes rather limited (Excell, 2000). However, textbooks that contain large numbers of multiple-choice questions are available in engineering (e.g. Floyd, 2007), optometry (e.g. Fletcher and Oliver, 1996), physiology (e.g. Colbert, 1996), mathematics (e.g. Bolt and Reynolds, 1978) and physics (e.g. Porter, 1987), among other disciplines.

These advantages and disadvantages suggest that multiple-choice questions should be used as one strand in a balanced and creative summative assessment regime.

2. Assessment details

Multiple-choice questions are used in formative and/or summative assessment mode by the author in a number of Level 8 and Level 9 programmes and modules. The details are summarised in Table 1.

Programme	Level and stage	Year	Module name	Formative and/or summative	n
DT235 BSc in Medical Physics and Bioengineering	Level 8, Year 3	2010	Feedback and Control	Both	8
DT021 BE in Electrical/Electronic Engineering	Level 8, Year 4	2009	Control Engineering	Summative	17
		2010		Both	11
		2009	Time Delay Systems	Summative	5
DT015 MSc in Energy Management	Level 9	2009	Energy Control Systems	Summative	14
DT702 ME in Pharmaceutical Process Control and Automation	Level 9	2010	Process Control Engineering	Both	9
			Advanced Control Engineering	Both	9

Table 1: Details of Level 8 and Level 9 programmes and modules which use multiple-choice questions (n = number of students).

When formative assessment is used, students submit answers to short, paper-based, multiple-choice quizzes. The quizzes are typically given at the end of each topic explored in the classroom. Students complete the quizzes anonymously, with the author communicating that the purpose of the work is to identify "muddy points" in student understanding. The solutions of the quizzes are then explored in the next classroom session, before a new topic is started.

In summative assessment, one question is made up of multiple-choice parts in the closed-book terminal examination, with the exception of the Control Engineering module. When used in the terminal examination, the question with multiple-choice parts is either compulsory (Time Delay Systems, Energy Control Systems, Process Control Engineering) or elective (Feedback and Control, Advanced Control Engineering), depending on the module learning outcomes. On the Control Engineering module, which is completely continuously assessed, open-book multiple-choice examinations are used in weeks 2 and 13 of the module. In all summative assessments with multiple-choice questions, negative marking is used if the solution offered is incorrect.

Sections 3 and 4 detail some relevant assessment experiences in the Control Engineering module, and all other modules, respectively.

3. Control Engineering module (Level 8, Year 4)

In the 2009-10 academic year, the subject was summatively assessed as follows:

- 75-minute open-book multiple-choice examination (Week 2). 20% of module mark. 50 questions, 4 answers.

- PowerPoint presentation on a control engineering topic (Week 7). Peer and tutor assessed; rubric available. 20% of module portion mark.
- Open-book design exercise using interactive learning modules. 30% of module portion mark (Week 10). A sample answer was used by the lecturer to assist in marking. A voluntary self-assessment exercise followed this, based on the sample answer.
- 90-minute open book multiple-choice examination. 30% of module portion mark (Week 13). 50 questions, 4 answers.

The following methods of formative assessment were used in addition:

- Sample 75-minute open-book multiple-choice examination (Week 1).
- Three multiple choice quizzes.

The purpose of the multiple-choice examination in Week 2 is to assess the pre-requisite knowledge for the module, as recommended by Felder and Brent (2001), among others. This was considered necessary, as students successfully completed two previous modules in the subject, the latest nine months previously (with the intervening time being devoted to a full-time work placement). In the first lecture of this module, the author introduced himself, and discussed, among other issues, learning outcomes for the present module and the manner in which the module would be assessed. Then, the author discussed the continuous learning approach to be adopted in the module, stating that module content builds on previous work and reminding the students of the learning outcomes for the previous two control modules. Chapters 2 to 19 of Wilkie *et al*. (2002) were referenced, with the author suggesting that the multiple-choice questions would be similar to a selection of those at the end of each chapter in this book; a sample multiple-choice examination was also provided. The author suggested to the students that it may be sensible for them to form study groups to prepare for the assessment. The assessment philosophy on the module is an open-book one, and this was used in the assessment; it was felt that this also eased the pressure on students, bearing in mind the assessment timescale.

One advantage of using multiple-choice questions for assessing pre-requisite knowledge is that the instructor can, by analysing the percentage of students correctly answering each individual question, assess the topics that students find most difficult, and treat these topics subsequently. This was done by the author.

The open-book multiple-choice examination in Week 13 followed the same structure as that in Week 2. When student assessment data is analysed for the 2008-9 and 2009-10 academic years ($n = 28$), a number of conclusions emerge:

- There is a borderline statistically significant relationship between performance in the open-book multiple-choice examinations and performance in the other assessments ($p = 0.0348$, correlation coefficient = 0.40);
- There is a statistically significant relationship between performance in the open-book multiple-choice examination in Week 2 and performance in the open-book examination in Week 13 ($p = 0.0096$, correlation coefficient = 0.48);
- There is a highly statistically significant relationship between the overall performance in the subject and performance in Automation, a related subject ($p < 0.00005$, correlation coefficient = 0.66). However, a student who scores 57 in Automation could expect to score 40 in Control Engineering, on average, indicating that students find the (latter) subject difficult.

The author also obtained student feedback on the use of multiple-choice tests, through the standard DIT student survey questionnaire, discussion with the class representatives, and, in 2009-10, from a student focus group. Introduced in 2009-10, the formative multiple-choice quizzes were generally considered a good idea by

students. The open-book nature of the assessment was generally considered fair, though concern was expressed about negative marking in the examinations. When the latter point was explored further, a number of students suggested that choosing the correct answer, with a penalty for choosing the wrong answer, was too demanding. Interestingly, when the author suggested that it was ethically important for engineering professionals not to propose solutions to problems in which they had less than full confidence, two students disagreed based on their work placement experience. They suggested that accuracy was not demanded of them, as their work was subsequently checked for errors. The author intends to discuss this issue in the first lecture with the full cohort of students in the next academic year. In addition, the time for the first multiple-choice examination will be increased to 90 minutes, to address concerns raised that 75 minutes was not sufficient.

4. Other modules

As detailed in Table 1, the author has used multiple-choice questions in formative and/or summative assessment in a variety of other Level 8 and Level 9 modules. However, only outline discussion is appropriate, as the number of students who have taken these modules is small.

- For Feedback and Control (DT235 BSc in Medical Physics and Bioengineering, Level 8, Year 3), 4 of 8 students chose to do the elective multiple-choice examination question. Three formative multiple-choice quizzes were used throughout the module. One student commented on the student survey questionnaire that a good feature of the module was that "constant feedback on performance allows for early improvements to be made to increase marks".
- For Energy Control Systems (DT015 MSc in Energy Management, Level 9), there is a statistically significant relationship between performance in the 'conventional' elective examination questions and performance in the compulsory multiple-choice questions ($p = 0.007$, correlation coefficient = 0.71, $n = 13$). Students did not give feedback on the use of multiple-choice questions.
- For Process Control Engineering (DT702 ME in Pharmaceutical Process Control and Automation, Level 9), there is a statistically significant relationship between performance in the 'conventional' elective examination questions and performance in the compulsory multiple-choice questions ($p = 0.006$, correlation coefficient = 0.83, $n = 9$). Ten formative multiple-choice quizzes were used throughout the module. In student comments on the module gathered through the student survey questionnaire, two students commented that "the constant feedback and communication of progress" through these quizzes was a good feature of the module.
- For Advanced Control Engineering (DT702 ME in Pharmaceutical Process Control and Automation, Level 9), there is not a statistically significant relationship between performance in the 'conventional' examination questions (one of which is compulsory) and performance in the elective multiple-choice questions. One formative multiple-choice quiz was used in the module. In a comment written in his examination answer book, one student stated that the multiple choice question "should have given me a better chance to show what I know, rather than what I am not absolutely sure of".

5. Conclusions

In conclusion, the author agrees that multiple-choice questions, in both formative and summative assessment mode, have the advantages of facilitating comprehensive student learning, and allowing analysis of student understanding, as detailed previously. The small number of students in the author's modules means that assessment efficiency is not a priority; it also means that it is difficult, in some cases, to show statistically significant relationships between different types of assessment outcomes. The author suggests that many Level 8 and Level 9 programmes in engineering have learning outcomes that make multiple-choice questions suitable as an assessment strategy (as part of a suite of assessment options). Students are more favourable to the use of multiple-choice questions in formative assessment, with negative marking of these questions in summative assessment attracting adverse comment. The author intends to continue to explore the use of multiple-choice questions in the modules for which he has academic responsibility.

References

Azalov, P., Azaloff, S. and Zlatarova, F. (2004). "Work in progress – comparing assessment tools in computer science education: empirical analysis", Proc. *34th ASEE/IEEE Frontiers in Education Conf.*, F2G-18 to F2G-19.

Azer, S.A. (2003). "Assessment in a problem-based learning course: twelve tips for constructing multiple choice questions that test students' cognitive skills", *Biochemistry and Molecular Biology Education*, 31(6), 428-434.

Bolt, R.L. and Reynolds, C. (1978). "Multiple choice questions in mathematics", Edward Arnold.

Brown, R.W. (2001). "Multi-choice versus descriptive examinations", Proc. *31st ASEE/IEEE Frontiers in Education Conf.*, T3A-13 to T3A-18.

Chang, S.-H., Lin, P.-C. and Lin, Z.-C. (2007). "Measures of partial knowledge and unexpected responses in multiple-choice tests", *Educational Technology and Society*, 10(4), 95-109.

Colbert, D. (1996). "Multiple choice questions in basic and clinical physiology", Oxford University Press.

Davies, T. (1994). "Multiple choice questions in Electronics and Electrical Engineering", Butterworth-Heinemann.

DeSantis, M. and McKean, T.A. (2003). "Efficient validation of teaching and learning using multiple-choice exams", *Advances in Physiology Education*, 17(1), 3-14.

DiBattista, D., Mitterer, J.O. and Gosse, L. (2004). "Acceptance by undergraduates of the immedicate feedback assessment technique for multiple-choice testing*", Teaching in Higher Education*, 9(1), 17-28.

Excell, P.S. (2000). "Experiments in the use of multiple-choice examinations for electromagnetics-related topics", *IEEE Transactions on Education*, 43(3), 250-256.

Felder, R.M. and Brent, R. (2001). "FAQs. IV. Dealing with student background deficiencies and low student motivation", *Chemical Engineering Education*, 35(4), 266-267.

Fenna, D.S. (2004). "Assessment of foundation knowledge: are students confident in their ability ?", *European Journal of Engineering Education*, 29(2), 307-312.

Fletcher, R. and Oliver, K.M. (1996). "Multiple-choice questions in optometry", Butterworth-Heinemann.

Floyd, T.L. (2007). "Principles of electric circuits", Pearson, 8th Edition.

McElroy, P., Brabazon, D. and McLoughlin, E. (2006). "Analysis of use of multiple choice question (MCQ) examinations as an assessment tool for postgraduate engineering modules", Proceedings of the *Irish Learning Technology Association Conference*, IT Sligo.

Multiple Choice and Matching Tests (2005). School of Graduate Studies and Continuing Education, University of Wisconsin. Available at http://www.uww.edu/Learn/multiple_choice_and_matching.php [accessed 27 May 2010].

O'Loughlin, E. and Osterlind, S. (2007). "A study of blended assessment techniques in on-line learning", Proceedings of the *AISHE Conference*. Available at http://www.aishe.org/events/2006-2007/conf2007/proceedings/paper-25.doc [accessed 28 May 2010].

Porter, G. (1987). "Senior physics: multiple choice questions", Folens.

Struyven, K., Dochy, F., Janssens, S., Schelfhout, W., and Gielen, S. (2006). "The overall effects of end-of-course assessment on student performance: a comparison of multiple choice testing, peer assessment, case-based assessment and portfolio assessment", *Studies in Educational Evaluation*, 32, 202-222.

Wilkie, J., Johnson, M. and Katebi, R. (2002). *Control Engineering: an introductory course*, Palgrave, U.K.

Zhao, Y.Y. (2005). "Algorithms for converting raw scores of multiple-choice question tests to conventional percentage marks", *International Journal of Engineering Education*, 21(6), 1189-1194.

Zhao, Y.Y. (2006). "How to design and interpret a multiple-choice-question test: a probabilistic approach", *International Journal of Engineering Education*, 22(6), 1281-1286.