

THE RELEVANCE OF MULTIPLE-CHOICE ASSESSMENT IN LARGE COHORT BUSINESS LAW UNITS

JOHN SELBY,* *PATRICIA BLAZEY*** AND *MICHAEL QUILTER****

This is the first part of a multi-part paper which discusses particular issues relating to the successful assessment of large student cohorts in Business Law subjects. This paper discusses assessment concepts, the relevance and value of multiple-choice testing in large student cohorts and the motivations of modern students. It will also assess the costs and benefits of multiple-choice assessment from the perspective of the administration of the Business Law Department at Macquarie University.

Future parts of this paper will discuss statistical techniques which can determine the validity, reliability and effectiveness of multiple choice as an assessment technique before presenting an analysis of the outcome of introducing multiple-choice assessment into core subjects (each with enrolments >800 students) within the Department of Business Law at Macquarie University. Those parts will highlight the evolution of our assessment policy to reconcile two previously conflicting goals: 1) enhancing deep student learning; and 2) managing assessment within workload constraints.

I. INTRODUCTION

The Department of Business Law at Macquarie University is experimenting with introducing multiple-choice assessment into its core subjects. This is the first of a multi-part paper which will detail the Department's experiences throughout this experiment. We have set out below the rationale for this experiment and the pedagogical basis upon which it is constructed.

II. LEARNING AND TEACHING CHALLENGES IN BUSINESS LAW SUBJECTS

In delivering its course offerings, the Department of Business Law at Macquarie University faces several significant learning and teaching challenges. First, the core subjects have very large enrolments, usually 800-1200 students per subject per semester.¹ Second, as a consequence of the Australian tertiary education system's success in attracting enrolments from outside Australia, over 60 per cent of those enrolments are international students coming from diverse backgrounds such as China, India, Indonesia, Thailand and Korea. Those students find that the common law is alien to the legal system established in their home countries. Also, English is often their second language and many do not have sufficient knowledge of English to deal with complex grammar and general vocabulary, let alone legal vocabulary.

Third, as a service department, the Department of Business Law's subjects are taught to students enrolled in at least 17 different undergraduate degree programs. Consequently, its subjects are ancillary to the majority of subjects the students study as they pursue their various degrees (the majority of enrolments are by students studying for commerce or business degrees).

Fourth, as a consequence of voluntary student unionism and increased financial pressures on students at the tertiary level, educators must provide interest and focus in an environment where universities have lost their role as social hubs. Increasingly, students spend less time at university and this translates to less time preparing and studying. For

* Lecturer, Department of Business Law, Macquarie University; Research Associate, Centre for Cyberspace Law & Policy, University of New South Wales.

** Head of Department and Senior Lecturer, Department of Business Law, Macquarie University.

*** Senior Lecturer, Department of Business Law, Macquarie University.

¹ The core subjects are Business Law, Corporations Law and Revenue Law.

many of our students, a university education is only one of several significant goals being simultaneously pursued.

Fifth, the ‘classic’ problems associated with large-group teaching remain. ‘At the risk of being oversimplistic, many of the standard complaints about the large lecture course can be traced to the impersonal nature of the classroom experience. “Ineffective,” “cold,” “distant,” and even “boring” — each of these descriptions can be linked in some way to an impersonal communicative environment’.²

Whilst the challenge of delivering interesting classes and engaging students remains a focal point in learning-and-teaching planning, there is an increasing need to consider and enhance the manner of student assessment. In a simple sense, if the student cohort has their focus upon completion of their degree and their professional career, as many of those doing accounting and similar professionally based degrees do, then assessment and assessment procedure are most likely the parts of their university experience that carry most meaning.

III. CHANGING ATTITUDES TOWARDS ASSESSMENT IN LAW

Andrew Ward and Alan Jenkins describe the change in assessment as occurring in different eras.³ In the 1970s, unseen exams were dominant, forming 100 per cent of assessment in law subjects. By the end of the 1980s, lecturers introduced coursework assessment in the form of essays, class participation, and group and single presentations.⁴

Those traditional forms of assessment within law have a number of drawbacks when applied to large student-cohort subjects. First, problem-solving or essay-based assessments can assess the depth of student learning but neither scale efficiently to assess the breadth of that learning.⁵ Whilst it is important to recognise the value of assessing students’ writing ability, it is perhaps less necessary to do so multiple times within one exam.⁶

Second, with 800-1200 students enrolled in a subject, it is impossible to have a single academic mark all of the students’ assessments. When more than one academic marks the same assessment, there is a challenge in ensuring that those academics consistently mark with equal difficulty levels. Preparing detailed, prescriptive marking guides can ameliorate this problem, but not remove it.

Third, the time and resources which must be applied to assessing the students’ output is very significant. The intensity of marking such high volumes of papers within a short period of time often leaves staff physically, emotionally and psychologically drained to the point where their research productivity is reduced during all-too-scarce non-teaching periods.

It is now recognised that no single form of assessment should comprise 100 per cent of the assessment types within a broader course of study.⁷ Although multiple-choice assessments have been used in many other areas of academia outside law schools for several decades, professional disciplines seem to have been the most resistant to introducing this form of assessment, although some medical schools have adopted it with the goal of assessing higher-order learning.⁸ There is evidence that some law schools have over 60 years history of using multiple choice as a component of their assessments. For

2 James Davis, ‘Dialogue, Monologue and Soliloquy in the Large Lecture Class’ (2007) 19(2) *International Journal of Teaching and Learning in Higher Education* 178, 178.

3 Andrew Ward and Alan Jenkins, ‘The Problems of Learning and Teaching in Large Classes’ in Graham Gibbs and Alan Jenkins, (eds) *Teaching Large Classes in Higher Education* (1992).

4 Ibid 33.

5 Lynn Daggett, ‘All of the Above: Computerized Exam Scoring of Multiple Choice Items’ (2007) 57 *Journal of Legal Education* 391, 392 and Vaughn Ball, ‘Objective Questions in Law Examination’s, (1959) 12 *Journal of Legal Education* 567.

6 Ball, above n 5, 568.

7 Daggett, above n 5, 392.

8 For the US experience, see Kathryn Knecht, ‘Assessing Cognitive Skills of Pharmacy Students in a Biomedical Sciences Module Using a Classification of Multiple-Choice Item Categories According to Bloom’s Taxonomy’ (2001) Winter 65 *American Journal of Pharmaceutical Education* 324. For the UK experience, see Paul McCoubrie, ‘Improving the Fairness of Multiple-Choice Questions: A Literature Review’ (2004) 26(8) *Medical Teacher* 709-712.

example, the first edition of the *American Journal of Legal Education* from 1948 included a comment article titled ‘The Validity of Objective Examination in Constitutional Law’, which discussed the experiences of academics when introducing multiple-choice assessment at the Ohio State University Law School.⁹ Douglas Miller even argues that ‘the validity of multiple-choice testing in the law field has been established by fairly careful study, and there appears to be fairly widespread acceptance of this form of testing not only on bar examinations, but also in law school examinations’.¹⁰ At least one UK law school has reported on its (positive) experience of using Bloom’s *Taxonomy* (discussed in greater detail below) when introducing multiple-choice assessments.¹¹

IV. INTEGRATING LEARNING THEORIES THROUGH TEACHING AND LEARNING INTO ASSESSMENT

There are many theories which attempt to describe how learning occurs, both in children and in adults.¹² Many of these theories overlap and many are contradictory. It seems that there is no overarching ‘general theory of learning’ which has gained broad acceptance. One concept that has been accepted for nearly 25 years is that students have the ability to combine both *surface* and *deep* approaches to learning.¹³ Nonetheless, student expectations of assessment are relevant to the approach they adopt and ‘the perceptions that students have of assessment even at the end of an exercise will dictate their approach’.¹⁴

Once a particular learning theory (or theories) has been selected for inclusion within a subject, it is necessary to consider not only how lectures and tutorials will incorporate the theory, but also the impact the theory will have upon both summative and formative assessments in that subject.

A. Bloom’s Taxonomy of Learning

Bloom’s *Taxonomy*¹⁵ is a famous structuring of educational objectives, first published in 1956, which has been selected by the Department of Business Law at Macquarie University for inclusion within its subjects. It is not the only taxonomy of learning, nor is it without its faults and criticisms.¹⁶ However, it is still recognised as a useful tool within academic scholarship on learning and has driven much research into attempts to increase deep learning amongst students.¹⁷

Bloom’s *Taxonomy* contains six major classes of educational objectives:¹⁸

- Knowledge;
- Comprehension;
- Application;

9 Paul Spayde and Frank Strong and W.R. Flesher, ‘The Validity of an Objective Examination in Constitutional Law’ (1949) 1 *Journal of Legal Education* 251, 252-256.

10 Douglas Miller, ‘Using Examinations in First-Year Legal Research, Writing and Reasoning Courses’ (1997) 3 *Legal Writing: Journal of the Legal Writing Institute* 217, 232-233. See also Linda Crane, ‘Grading Law School Examinations: Making a Case for Objective Exams to Cure what Ails “Objectified” Exams’ (2000) 34 *New England Law Review* 785, 804-8.

11 Edwina Higgins and Laura Tatham, ‘Exploring the Potential of Multiple-Choice Questions in Assessment’ 2(1) *Learning and Teaching in Action* 1, 6.

12 See, eg, the lengthy list of learning theories at emTech, *Learning Theories* <http://www.emtech.net/learning_theories.htm> at 4 December 2008.

13 Paul Ramsden, ‘The Context of Learning’ in Ference Marton, Dai Housnell and Noel Entwistle (eds), *The Experience of Learning* (1984).

14 John Heywood, *Assessment in Higher Education* (2000) 225.

15 Benjamin S Bloom (ed), *Taxonomy of Educational Objectives: The Classification of Educational Goals, Handbook 1: Cognitive Domain* (‘Bloom’s Taxonomy’).

16 See generally: Lorin Anderson and Lauren Sosniak (eds), *Bloom’s Taxonomy: A Forty-Year Retrospective* (1994). The three most common criticisms of the taxonomy are: the distinctions between any two levels may be blurred; it is not hierarchical, only a set of categories; and the lockstep sequence underlying the taxonomy based upon ‘complexity or difficulty’ is naïve — Neville Postlethwaite, ‘Validity vs Utility: Personal Experiences with the Taxonomy’ in Lorin Anderson and Lauren Sosniak, *Bloom’s Taxonomy: A Forty-Year Retrospective* (1994) 175.

17 Robert Wood, ‘Multiple Choice: A State of the Art Report’ (1977) 1 *Evaluation in Education* 191, 204.

18 Bloom, above n 15, 18.

- Analysis;
- Synthesis; and
- Evaluation.

Knowledge involves 'the recollection of an idea, material or phenomenon in a form very close to that in which it was originally encountered'.¹⁹ Comprehension involves 'objectives, behaviours or responses which represent an understanding of the literal message contained in a communication', and includes translation, interpretation and extrapolation.²⁰

Application requires students to demonstrate their ability to apply comprehended knowledge to previously unseen types of problems. Reaching a solution to the problem may also require the student to use the other stages of the taxonomy.²¹ Analysis requires students to be able to demonstrate their ability to 'breakdown material into constituent parts and then detect relationships between those parts and the way they are organised'. The border between comprehension and analysis is recognised as being porous.²²

Synthesis requires students to take separate components and build them together into a coherent whole.²³ Evaluation requires the application of criteria and standards in a qualitative or quantitative way so as to 'make judgments about the value or purpose of ideas, works, solutions, methods, material, etc.'²⁴

B. Why was Bloom's Taxonomy Selected?

Adopting Bloom's *Taxonomy* offered a number of benefits. First, it has been widely used and tested in higher education.²⁵ Second, there is a large body of literature showing strengths and pitfalls when applying the Taxonomy. Third, it is a relatively practical and implementable tool. Fourth, it is structured in a way which the Department believes can be applied to aid our students' learning. Fifth, and finally, it has been applied by a number of other (mainly US-based) law schools when introducing multiple-choice assessments and, therefore, it is appropriate to be able to compare like with like.²⁶

V. DESIGNING EFFECTIVE ASSESSMENT FOR BUSINESS LAW SUBJECTS

Effective assessment is a complex mix of matters incorporating characteristics of validity, reliability and utility. The dichotomy between measurement and judgment is sometimes also the dichotomy between short-form 'efficient' assessment and time-heavy problem-based assessment. However, this is not necessarily irreversible and 'the issue of whether achievements can be measured surfaces most strongly where the expectations of students include aspects of performance as a beginning professional'.²⁷

Students in Business Law subjects are clearly 'beginning professionals' and the task to design an effective measuring process is at the core of any successful large offering. From an academic's perspective, assessment outcomes can involve many goals, but to students it is their result that is primary. This would suggest that summative assessment characterised by high validity and reliability would be essential to meet their career-focused needs. However, some level of foundation to enable them to meet their future challenges, introduced by way of formative assessment, can still play a part.

19 Anderson and Sosniak, above n 16, 16-18.

20 Ibid 19-20.

21 Ibid 20-21.

22 Ibid 21-23.

23 Ibid 23-24

24 Ibid 24.

25 For an example of research similar to this project, but from a medical school perspective, see Knecht, above n 8, 324.

26 See, eg, Daggett, above n 5, at p408 footnote 49, where she states 'I generally try to write items that test at the analysis or higher levels of Bloom's Taxonomy ... rather than items which merely ask students to recall concepts or apply them in a simplified manner'. See also Greg Sergienko, 'New Modes of Assessment', (2001) 38 *San Diego Law Review* 463, 495, where he states 'Using one of the systematic descriptions of intellectual skill levels, such as Bloom's Taxonomy ... helps ensure that questions are drafted to test an appropriate range of skills'.

27 Mantz Yorke, *Grading Student Achievement in Higher Education: Signals and Shortcomings* (2008) 26.

Assessments that take place within modules may be formative, providing feedback on performance, and summative, in that they count towards the grade to be awarded for performance on the module as a whole. Examples of such assessments are multiple-choice tests.²⁸

In fact, with the highly specialised ability of information technology to assess and break down multiple-choice results, they have developed into a particularly potent diagnostic tool of overall standards and trends in specific groups of students.

What then is effective assessment? Sally Brown and Peter Knight²⁹ include clarity of purpose, the ability to review progress by all parties, assessing what is claimed to be assessed, being credible, cost-efficient, having clear outcomes and being subject to quality assurance. A particular factor, if successful assessment is to be maintained, is for assessment to produce usable data. If the data obtained is to fulfil its potential and provide valid performance indicators about the core business of higher education, that is teaching and learning,³⁰ an essential characteristic of its collection must be its reliability.

In commenting on the results of an earlier study on reliability of assessment, Mantz Yorke stated that ‘the level of consistency was reduced where there was greater latitude for interpretation on the part of the assessors as regards whether the criterion had been met’.³¹ In disciplines such as law, responses to problem-style questions vary widely, making parity of assessment difficult. The basic content, the law from which the answer should be drawn, is generally the same amongst examinees. However, writing styles, structure and perceptions of relevance underpin the variance. This creates substantial difficulties in reliability and consistency between examiners and the problem is exacerbated where the student population has differing levels of English language proficiency. There are often errors of fact and law in such pieces of work. Plagiarism is a growing problem and lecturers spend hours on the internet or using software tools to seek out such material.³²

A. Assessing Deep Learning Through Multiple-Choice Assessment

An obvious criticism of multiple-choice examinations is that in their highly summative role they encourage *surface* rather than *deep* learning. However, many factors mitigate this narrow view of multiple-choice examinations. Firstly, not all multiple-choice examinations are the same. There is ample scope for questions to be complex and testing of various levels of content. Secondly, in law units, where students generally receive tutorials revolving around problem-solving, it is acceptable for multiple-choice questions to take similar problem formats requiring examinees to work through a set of facts before being able to approach the weighing of the alternatives provided. Where the examinees are largely non-lawyers doing law units, such as Business Law students, identifying issues and postulating outcomes is perhaps more important than the ability to coherently express such outcomes in written form. Clearly, though, students going on to practise law need this ability. Further, multiple-choice testing will not generally be the sole determinant of a student’s grade and a student will be exposed to a wider group of educational skills throughout their higher education. Diversity in assessment is almost universally accepted and multiple-choice examinations fit into, rather than dominate, this diversity.

Multiple-choice examinations do not have to stifle a *deep* approach merely because they are efficient to undertake and to grade. In fact, summative examinations such as multiple-choice ‘can have considerable power to encourage learning, partly through providing that extrinsic motivation which is so necessary ... [and] can provide a stimulus

28 Ibid 12.

29 Sally Brown and Peter Knight, *Assessing Learners in Higher Education* (1994) 23.

30 Ibid 139.

31 Yorke, above n 27, 58. The earlier study was David Baume and Mantz Yorke, ‘The Reliability of Assessment by Portfolio on a Course to Develop and Accredite Teachers in Higher Education’ (2002) 27(1) *Studies in Higher Education* 7.

32 Search engines such as Google <<http://www.google.com>> at 4 December 2008, and anti-plagiarism software like TurnItIn <<http://www.turnitin.com>> at December 4 2008, are commonly used to find evidence of plagiarism.

for understanding to be developed through *deep learning*'.³³ Whereas the format for response varies between written examinations and multiple-choice examinations, the desired attributes, which include a grasp of content, application, interpretation and comprehension, do not. A criticism of multiple-choice examinations is that they encourage study without reflection and an example of this can be seen when students guess answers. However, this is merely an instance of poor strategy which similarly exists in all examination formats. The most common characteristic of poor responses in written examinations is not answering the question — this, like guessing in multiple-choice examinations, is simply an inappropriate approach to the task. Even detractors of the multiple-choice format concede that 'when multiple-choice questions are used in the more formative ways they have great potential as a learning and teaching tool'.³⁴

Multiple-choice examinations should not be seen as substitutes for those parts of student assessment requiring expression such as presentations and class performance. They are, however, alternatives to traditional problem-solving formats where essay-style answers are required. An inherent difficulty with written response examinations is that because a student's time during the examination is largely taken with writing, the scope or coverage of the examination is limited. Further, student perceptions of the need to digest large amounts of material (because they have an expectation that large amounts must be disgorged in their answers) results in overloading, sometimes confusion, and often cramming. Multiple-choice examinations on the other hand can cover unit content widely without the need to emphasise or specialise. This can counter cramming as the foundation of the multiple-choice examination relates to the overall delivery of the unit and the overall grasp of the material. Accordingly, because multiple-choice examinations involve a large number of discrete questions, they not only can directly assess intended learning outcomes of the whole unit but, importantly, 'many of the individual learning outcomes can be addressed in a single assessment'.³⁵

Enhancing student perceptions of their tasks in examinations is an important aspect of worthwhile assessment practice. The best assessment strategy can fail if it does not take account of how students see and approach examinations. Accordingly, successful multiple-choice testing of large groups of students requires educating academics about new approaches for assessment. Multiple-choice examinations can be complex and demanding, and issues can be raised requiring several levels of content. It is not really a question of 'choice' at all sometimes but a question of 'exclusion'. In the past, the perception of multiple-choice testing as addressing *surface* learning has meant that the format has been regarded as superficial. However, that is not necessarily the case and

'multiple-choice questions need not be limited merely to testing knowledge; tests can include more challenging comprehension/application-based questions ... and it is only if they are inappropriately used or poorly designed that there is a risk of 'dumbing-down'.³⁶

It is even suggested that well-designed multiple-choice examinations are useful in distinguishing the *surface* learner from the *deep* learner.³⁷ This will be achieved if analysis and evaluation are built into the assessment.

Although successful multiple-choice questions are often difficult to draft there, is a substantial benefit in time saved by freeing academics from marking. This is imperative in large undergraduate units and results in a value-based assessment process. The corollary of this is time to attend to students. In large units, finding ways to assess students that save academic time are positive factors in student welfare. Provided multiple-choice assessment is predictable, consistent, relevant and equal for all students, it serves a summative, and

33 Brown and Knight, above n 29, 67-68.

34 Moragh Paxton, 'A Linguistic Perspective on Multiple Choice Questioning' (2000) 25(2) *Assessment and Evaluation in Higher Education* 109. Paxton is critical of multiple-choice testing in economics units.

35 Sue Bloxham and Pete Boyd, *Developing Effective Assessment in Higher Education: A Practical Guide* (2007) 212.

36 Higgins and Tatham, above n 11, 1.

37 *Ibid.*

possibly a formative, purpose. If thereby academics have more time to advise and counsel the large number of students' dependent on them, the process of education is enhanced.

Another advantage of multiple-choice assessments is that feedback to students can be given much faster than traditional forms of assessment. Robert Wood³⁸, Greg Sergienko³⁹, Steven Friedland⁴⁰ and Lynn Daggett⁴¹ all note the widespread survey evidence that prompt feedback has a significant impact on increasing student learning, especially in helping students to self-identify and rectify under-performance midway through the semester.⁴²

B. Incorporating Bloom's Taxonomy When Designing Multiple-Choice Assessments?

Designing multiple-choice exams is much harder than creating a problem-based combination essay examination. This is because multiple layers of difficulty and analysis can and need to be incorporated into questions.⁴³ The level of difficulty will depend on the type of student, the level they are at and the subject itself.

There are two different assessment goals for which multiple-choice assessments can be used: 1) Norm-referenced evaluation; and 2) Criterion-referenced evaluation. The design and types of questions used within a multiple-choice assessment will vary depending upon which goal is to be achieved.

Norm-referenced evaluation seeks to distinguish between high-performing and low-performing students (in terms of their mastery of the subject area) and, therefore, a multiple-choice assessment designed around this goal would contain questions which have various levels of item difficulty and which are highly statistically discriminatory.⁴⁴

Criterion-referenced evaluations seek to assess whether students have mastered specific concepts and skills⁴⁵ (eg, pass/fail) and, therefore, the questions will be less likely to be statistically discriminatory and fewer questions would be at the highest or lowest difficulty levels.

Both norm-referenced and criterion-referenced assessments should be based upon the learning objectives of the subject.⁴⁶ Edwina Higgins and Laura Tatham recommend a strategy of providing advance notice to students of sample multiple-choice questions or formative quizzes to overcome any student aversion to that form being used in summative assessment.⁴⁷

C. Assessing Knowledge through Multiple-Choice

Possessing knowledge is an important precursor to higher levels of learning and, therefore, it is still appropriate to ask knowledge-focused multiple-choice questions on an assessment. However, knowledge questions should only form a minority of the total questions asked.⁴⁸

Knowledge questions should not attempt to assess obscure trivia, for example, 'In what year was the Civil Liability Act passed in NSW?' Questions should instead focus on

38 Wood, above n 17, 201.

39 Sergienko, above n 26, 471.

40 Steven Friedland, 'A Critical Inquiry into the Traditional Uses of Law School Evaluation' (2002) 23 *Pace Law Review* 147, 200-201.

41 Daggett, above n 5, 398.

42 Sergienko, above n 26, 486.

43 Kenney Hegland, 'On Essay Exams' (2006) 56(1) *Journal of Legal Education*, 140.

44 Daggett, above n 5, 401.

45 Ibid.

46 Gerry Hess, *Exams and Other Methods of Assessment* (2001) Association of American Law Schools <<http://www.aals.org/profdev/nlt2001/hess.html>> at 4 December 2008.

47 Higgins and Tatham, above n 11, 4-5.

48 John Carneson, Georges Delpierre and Ken Masters, *Designing and Managing Multiple Choice Questions: Appendix C: MCQs and Bloom's Taxonomy* (1996) University of Cape Town <<http://web.uct.ac.za/projects/cbe/mcqman/mcqappc.html>> at 4 December 2008.

demonstrating knowledge of, for example, the most important case in a specific area of law; the meaning of a legal term; or ‘Which of the following legal propositions is false?’⁴⁹

D. Assessing Deep Learning Beyond Knowledge-focused Multiple-Choice Questions

1. Assessing Comprehension

This type of multiple-choice question is designed to assess each student’s ability to grasp the meaning of written or visual material (eg, a court hierarchy). Creative examples of this type of question include: showing a timeline of a significant/fictional case as it progressed through the appeal system and asking students to identify certain aspects of that process, or giving a short quote from a judgment and requiring students to identify which sentence was the ratio/obiter, etc.

2. Combination Questions

Another type of complex question, which can be used to test recognition of legal principles and the ability to identify the relevant legal principles in a factual scenario, presents students with a fact scenario and a series of potential legal principles. The students are first asked to identify which of those potential legal principles are correct legal principles. The follow-up question asks students to identify which of the principles are most relevant to the factual scenario,⁵⁰ assessing analytical skills.

3. Assessing Application

This is a fertile area to assess law student’s problem-solving abilities because questions can require students to apply a stated legal principle to a short fact scenario and to decide which conclusion would most likely be supported on those facts.

4. Assessing Analysis

Multiple-completion questions can be used to assess analysis skills. Such a question provides students with a short story and a series of related statements.⁵¹ The students are then required to select which of those statements can best achieve a certain outcome. In testing higher level skills in law, for example, such a question could contain a fact scenario and a series of additional facts from which the students must choose the combination which, if assumed, would most likely lead to an increased likelihood of a particular legal conclusion being reached. Similar structured questions can be used to test students’ ability to identify relevant legal issues, remedies available or applicable case law.⁵² Miller does caution that the item discrimination level of these types of questions does need to be monitored carefully.⁵³

5. Assessing Synthesis

Kathryn Knecht admitted finding it difficult to construct questions to assess synthesis amongst medical students, though her research showed success in the other five stages of Bloom’s *Taxonomy*.⁵⁴ The Centre for Instructional Technology at the Massachusetts Department of Higher Education argues that ‘by definition, synthesis cannot be assessed using multiple-choice questions’.⁵⁵

49 Knecht, above n 8, 326, 333: Knecht argues that ‘negative stem questions unnecessarily complicate test questions’ and increase the difficulty level of the question as compared to a positively-phrased but otherwise identical question. She argues that such questions should be restricted to situations where reinforcement of the need to avoid common mistakes is the goal.

50 Ball, above n 5, 570. It is worthwhile reviewing the helpful examples in the appendix to Ball’s article on pp573-577.

51 Wood, above n 17, 213-5.

52 Sergienko, above n 26, 492-3.

53 Miller, above n 10, 234-5.

54 Knecht, above n 5, 327.

55 Massachusetts Department of Higher Education, *Writing Multiple Choice Questions that Demand Critical Thinking* <<http://cit.necc.mass.edu/atlt/TestCritThink.htm>> at 4 December 2008.

6. Assessing Evaluation

Evaluation requires students to demonstrate the ability to judge the value of material for a given purpose.⁵⁶ One possible way to assess evaluation skills is to provide students with an answer to a short answer question from a previous year and to have the students evaluate the correctness, or identify the weaknesses, in that answer. Alternatives include having the students critique sample answers to problems, advice to partners in a law firm, pieces of legislation or topical newspaper articles discussing legal issues.

E. Challenges When Using Bloom's Taxonomy in Multiple-Choice Assessments

Using Bloom's *Taxonomy* to assess and revise the questions asked in multiple-choice assessments encourages academics to move beyond the common critique that such assessments 'only test recall of basic knowledge' and it is the goal of the Business Law Department to progressively increase the rate at which it assesses higher levels of learning in the taxonomy through multiple-choice assessment. The outcomes of such changes to the assessment process will be monitored and assessed.

Placing specific assessment tasks into the different classes within Bloom's *Taxonomy* requires academics to firstly be aware of the differences between those classes.⁵⁷ The challenge is to try to get consistent consensus about which class a particular assessment task falls within.

Lorin Anderson argues that asking only lower-order questions of students is an 'effective way of producing lower-order learning'. Consequently, it is important to structure multiple-choice examinations to ensure that both lower and higher level learning is assessed.⁵⁸ Unfortunately, surveys of examinations suggest that most items (80 per cent in some studies) are poorly designed and assess only the lowest level of the taxonomy, knowledge.⁵⁹ The difficulty of an assessment task is independent of the class within which it falls.⁶⁰

It is important not to simply increase the length and complexity of knowledge-based questions in an attempt to force students to engage in analysis, evaluation or synthesis. Complex wording in questions can be used to test some aspects of students' comprehension skills, but Wood argues that it is not the best way to assess higher level learning.⁶¹

VI. COSTS AND BENEFITS OF INTRODUCING MULTIPLE-CHOICE ASSESSMENTS INCORPORATING BLOOM'S *TAXONOMY* OF LEARNING INTO BUSINESS LAW SUBJECTS AT MACQUARIE UNIVERSITY

There are a number of benefits expected from the introduction of multiple-choice assessments incorporating Bloom's *Taxonomy* of learning into Business Law subjects at Macquarie University. First, greater breadth of student learning will be assessed. Second, student objections regarding inconsistency in marking between academics will be reduced through objective computerised marking.⁶² Third, the incorporation of software-based, customised student feedback systems will address one of the primary frustrations of recent student cohorts, namely lengthy delays between assessment pieces being sat and feedback being received. Fourth, a reduced emphasis (i.e. reduced from 90 per cent to 40-50 per cent) on writing traditional extended essay and problem-solving assessment tasks should

56 Carneson, Delpierre and Masters, above n 48, 2.

57 Amelia Kreitzer and George Madaus, 'Empirical Investigations of the Hierarchical Structure of the Taxonomy' in Lorin Anderson and Lauren Sosniak (eds), *Bloom's Taxonomy: A Forty-Year Retrospective* (1994) 68.

58 Lorin Anderson, 'Research on Teaching and Teacher Education' in Lorin Anderson and Lauren Sosniak (eds), *Bloom's Taxonomy: A Forty-Year Retrospective* (1994) 133. See also Harry Gensler, 'Valid Objective Test Construction' (1986) 60 *St. John's Law Review* 288, 290.

59 Anderson, above n 58, 138.

60 Kreitzer and Madaus, above n 57, 70.

61 Wood, above n 17, 210.

62 Yorke, above n 27, 22.

provide opportunities for students with various levels of English-language proficiency to display their abilities at both surface and deep levels. Fifth, academic staff stress levels during assessment periods should be reduced. Sixth, academic staff should have greater opportunities for productive research which will increase the overall research productivity of the department. Seventh, the department expects to reduce its expenditure on casual marking by tens of thousands of dollars per semester. Those funds can then be redeployed to hire more academics/research assistants and for conference travel. Finally, once a sufficiently large and validated bank of multiple-choice questions has been developed, such a bank may be licensed to Business Law departments at other Australian universities.

It must be recognised that there are also costs which will be incurred to achieve these benefits. The largest cost is the investment made by subject coordinators in learning how to incorporate Bloom's *Taxonomy* into their subjects and how to write effective multiple-choice assessments which test both surface and deep learning. The present workload model within the department arguably does recognise sufficiently the time required to build and maintain a sufficiently large item bank of validated multiple-choice questions suitable for use in assessments. It is also necessary for the department to invest in software suitable for constructing and statistically validating multiple-choice assessments (textbook publishers have provided some software to assist in this task) and for providing automated feedback to students. It is hoped that there may be scope for collaboration with Business Law departments at other Australian universities to defray and distribute some of these expenses.

VII. CONCLUSION

We have argued in this paper that there is a place for using multiple-choice assessment within the Business Law curriculum. However, such examinations are only capable of effective assessment when based on firm pedagogy. There are various benefits from introducing this form of assessment, ranging from increased reliability in marking, more efficient delivery of feedback to students and greater cost-effectiveness. It is important that academics fully understand the complexity and costs of introducing and validating multiple-choice assessments before pursuing this path.

Future parts of this paper will detail the techniques used by the Department of Business Law when constructing multiple-choice assessments and the statistical techniques used to assess reliability, validity and effectiveness of particular questions and entire exams, before reporting on the effectiveness of the introduction of this assessment form at Macquarie University.