
Student Assessment in MSOR

Penny Bidgood
Kingston University

bidgood@kingston.ac.uk



Bill Cox
Aston University

w.cox@aston.ac.uk



A teaching and learning
maths project funded by
LTSN Maths, Stats & OR
Network

As recent events attest, student assessment is a tricky business! In Quality Assessment and Subject Review of HE across the UK it has emerged as one of the most problematic areas [1] in MSOR provision. This article is a preliminary outline of the practitioners' views of good practice and areas for improvement in assessment, as revealed by the survey of subject review reports as part of the LTSN Maths, Stats & OR Network funded 'Where are we now?' project. We also mention some relevant material from mathematical education and staff and educational development, and suggest some areas that might benefit from further research and development. In effect this article forms a contents list for further material to come. The Network has recognised the importance of assessment and the need to disseminate good practice through its CAA series [2] and through the recent establishment of a project *Supporting good practice in assessment in MSOR* [3].

For discussion we will adopt a framework for assessment that broadly reflects the time sequence: setting and scheduling of assessment tasks; marking and moderation; feedback to students; use of outcomes of assessment.

SETTING AND SCHEDULING OF ASSESSMENT TASKS

The main issues that arise here are

- the link between what students are expected to be capable of (learning objectives), and the assessment tasks set
- match of assessment to student profile
- management and administration of the design of assessment
- setting and moderation of assessment tasks
- scheduling of assessment tasks

Link between learning objectives and assessment tasks set

This area, in which the assessment is linked to what the students are actually supposed to have learnt was commended in a quarter of providers, but criticised in the same number. It really relates to whether the assessment actually does the job it is supposed to – measure what students have learnt. Many providers now publish a complete list of explicit learning objectives and can relate these to the assessment tasks set in such a way that success in the latter demonstrates their achievement. This is particularly noteworthy in the case of transferable skills, where the link between skills and assessment is transparent, although few providers assess transferable skills directly. Assessment is not always well matched to learning objectives in the case of the higher order skills (synthesis, analysis, evaluation, etc). Another breakdown occurs when the assessment task becomes predictable, for example when examinations are similar from year to year or have a lot of bookwork content.

One third of providers were commended for the wide range of assessment instruments used to address the learning objectives. These include various sorts of written exams, coursework, projects and case studies, practical tests (especially in statistics). Innovative use of peer and self-assessment were also seen, and a number of providers now use presentations and poster sessions, portfolios, student logbooks. Specific recommendations for a wider range of assessment were made in some 10% of providers. Some recommendations for a wider range of assessment were made specifically in order to benefit weaker students – usually a recommendation for the extension of coursework. Coursework is usually limited to about 10-20%. The disparity between coursework marks and exam marks in the same module was sometimes a problem in this area.

How, in practice do we achieve a good match of assessment to learning objectives? Most assessment in MSOR in the UK is still conducted by formal, unseen, closed book exams. Most of us operate at a rather intuitive level in writing our exam questions, and we may need some practice in the framing and testing of learning objectives. To frame learning objectives we need a taxonomy of different types and levels [4,5,6 provide generic references in this area]. There has been some work recently on simplifying this for the MSOR practitioner. Smith et al [7] have given examples of how to write questions covering the full range of learning objectives in the eight category MATH hierarchy. Galbraith and Haines [8] used a three tier taxonomy: mechanical, interpretive, constructive and used it in question design. With only three pretty obvious categories, it is within the range of the average academic who is prepared to be a little more systematic about how they set exam questions.

Kahn and Kyle [9] contains much accessible discussion on the relation between learning objectives and assessment, from a wide range of perspectives. The Network holds this as one of its priorities and the workshop 'Assessment for a purpose' was devoted specifically to this area [10]. Matching assessment to learning was the subject of a Network workshop in May 2001 [13]. It also features significantly in the Benchmark statement for MSOR [14].

Match of assessment to student profile

About 20% were commended in this area, and the same number criticised. One symptom of a mismatch here is the 'long tail of weak students'. Another mismatch sometimes occurs when groups of students at different levels are taught together – eg BSc and MMath, or HND and BSc. Specific examples of good practice include the greater emphasis on coursework when incoming students are less well experienced in formal exams and differentiated assessment of HND and degree appropriate to the incoming students. Some providers were commended for the way in which assessment enabled students of all ranges of ability to receive appropriate reward for their efforts. Diagnostic assessment on entry is now almost essential, and reviewers commended it in about a quarter of providers. Good practice was found in the use of such assessment in easing the transition to university and in identifying the needs of students. The literature on diagnostic testing is extensive [9, 15]. The Network has just set up a project to survey diagnostic test provision [16]. The link between student profile and assessment is recognised in the Benchmark statement for MSOR [14].

Management and administration of the design of assessment

This area attracted a great deal of criticism (about 30% of providers). The general administration and management of the design of assessment is usually commented on by the external examiners, and usually reviewers were able to confirm their positive comments. Documentation and the clarity of assessment processes was however a major issue here. While a third of providers were commended in this, nearly 20% were criticised. The main point related to the students' lack of awareness or understanding of assessment processes. In cases of good practice assessment arrangements are well documented in student handbooks. Some providers give information to students on assessment on the university intranet or the web.

Setting and moderation of assessment tasks

About a fifth of providers were commended in this area, but the same proportion criticised. Positive comments about moderation of assessment tasks (as distinct from moderation of marking – see below) could be made in many cases, and there are many examples of good practice that will be reported elsewhere. The main negative comments related to the breakdown of moderation or the lack of evidence for it, even if it took place. There seems to have been little research on the efficacy, reliability and practical implementation of moderation.

Scheduling of assessment tasks

Roughly 15% were commended and a similar proportion criticised in this area. Modularisation has sometimes led to problems with examination scheduling. Also, fitting in such things as moderation means that a lot more needs to be done at each stage of the assessment process, putting more constraints on scheduling. So far as scheduling arrangements for coursework is concerned reviewers praised a number of providers for the care taken to avoid bunching of submissions. Comments were overwhelmingly favourable on timely feedback to students.

MARKING AND MODERATION OF ASSESSMENT

Issues that have emerged from the Subject Review reports are

- Marking schemes and assessment criteria
- Checking and double marking
- Reliability of marking
- Role of external examiners

Marking schemes and assessment criteria

Generally, the criteria for assessment were clear and understood by students, and good use was made of clear marking schemes. However, nearly 20% of providers were criticised in this respect. Assessment criteria were not always given or marking schemes not complete, or they are not used consistently, particularly in the case of projects. The relation between criteria and levels is not always clear in cases where the same course can be taken at different levels. The criteria applied to the use of raw marks in, for example, scaling or determination of degree classification, are not always made clear.

Checking and double marking

Most providers now have some system of checking and double marking of assessment, and this was often complimented for its effectiveness and rigour. However, over a third of providers had problems in this area. Sometimes the procedures for double marking were not clearly defined, or reviewers were disappointed by the extent of double marking. Sometimes only projects were double marked. The amount of double marking of exam scripts was very variable, ranging from none at all to all scripts. Again, this is an area where little research has been done and some sort of practitioner consensus would be useful to provide a balance between what is desirable, and what is practical.

Reliability of marking

Usually reviewers could report that marking was rigorous and reliable, but in 10% of cases some concerns were expressed relating to this area. Consistent marking of projects was sometimes singled out for praise. One or two institutions were commended for the way in which university or school oversight and systems helped to promote consistency of assessment marking and reporting. Overgenerous marking was found in some cases. Sometimes a disparity of marking between courses was found, with some failure rates as high as 50% observed. Reasons for such variability of marks between subjects in MSOR are advanced in the benchmark statement [14], by appealing to the special difficulties of mathematics, but no evidence is presented for this. Exams should be set and courses delivered so that scaling is not required, and mark distributions are reasonable.

Role of external examiners

With external examiners likely to play a more important role in quality issues, it is comforting to find that their role in MSOR is endorsed in most cases, and strongly so in at least a quarter. Besides protecting standards and students'

interests and commenting on exam processes generally, external examiners are sometimes involved throughout the year in course and curriculum development and other quality issues. Many different arrangements were found for the reporting processes used by external examiners. However, there was sometimes a wish expressed for greater involvement by external examiners, especially in respect of coursework and curriculum review. The information supplied to external examiners could sometimes be improved, as could the depth and rigour of some external examiners' reports.

Reviewers commended the response to external examiners in a third of cases, but deficiencies were found in 20%. The main criticism related to tardiness in dealing with external examiners' comments. Some providers were commended for the way in which they fed back to the external examiner on action taken.

FEEDBACK TO STUDENTS

While half of people were commended in this important area, nearly a third were criticised. Usually the comment was that the feedback was variable in quality and/or quantity to various degrees, and didn't always do much to support learning. Some reviewers expressed disappointment at the amount of feedback on students' scripts and held the view that verbal feedback in class alone was not adequate. Some providers use a standard feedback sheet. In general the quality and quantity of feedback was not always monitored adequately. However, there is a great deal of good practice in this area, which will be described elsewhere. There were a wide range of feedback mechanisms noted, including full solutions, annotations on scripts, oral feedback, classroom demonstrations, personal summary sheets, feedback forms, group sheets identifying common problems, posting on notice boards or the web, workbooks, follow-up teaching sessions, effective open-door policy, individual discussions with students, reworked exams.

USE OF OUTCOMES OF ASSESSMENT

Issues here are

- Progression and completion satisfactory?
- Distribution of awards satisfactory?
- Quality of student work
- Use of assessment results in MIS

Progression and completion satisfactory?

While a third of providers were commended for their student progression and completion, over 40% were criticised in some respect. Problems in progression and completion are manifested partly through outcomes of

assessment, and are a major problem in HE generally. Even in the top universities it has been noted that those students who progress may not always fulfil their potential (the 'long weak tail'). Generally, reviewers made one of two types of positive comment in this area – one where the failure/attrition rate is low in absolute terms, and the other when the completion may be low, but still impressive for the weak intake. Normally, it is the first and second year boundary at undergraduate level that attracts most criticisms about progression. A few providers already had measures in place to address such problems which included such things as flexible pathways, 'rescue' facilities, bridging courses, maths centres, remedial teaching, diagnostic tests, resits, etc. That many providers did not attract censure in this area shows that this problem can be solved. In terms of the nature of assessment, most difficulties seem to occur where students with non-standard entry qualifications are concerned.

There are a number of components to eventual completion - how long do students take to complete, do they leave with any exit qualification if they fail to complete the degree, what support mechanisms are in place to assist completion? Completion was commended in over a quarter of providers. Criticisms about the completion in normal time were quite widespread (nearly 20%), but in general there was little in terms of analysis of the profiles of students that failed to complete. Students often drop out for personal reasons or because of financial pressures, rather than through academic failure.

It is reassuring that a third of providers were actually commended for their added value provision. The main manifestation of added value was the extent to which students with poor or non-standard entry qualifications did just as well as more highly qualified students. Some providers with uniformly high intake qualifications were still able to demonstrate significant added value.

Distribution of awards satisfactory?

Related to progression, completion, and added value, this is dependent on the nature of the institution. There is little positive said about this area that can provide robust ideas about what to expect from degree classification profiles in general. External examiners usually confirm that the profile is broadly in line with similar institutions or courses elsewhere, and reports sometimes confirmed this. Sometimes reviewers commented that the profile was broadly in line with previous years, or the national average for MSOR (22 (First), 32.8 (2.1), 29.3 (2.2), 16.0 (Third/Pass) in 2000, [11]).

Not much work has been done on the spread of degree

awards across MSOR providers. One related piece of work is that by the SACWG group [11], which is a group of administrators and academics from nine educational institutions who collaborate in order to study assessment and related matters. The paper is devoted to investigating whether the mark-spread in degree classification data evident in HESA's data could be partly due at the module level to the method of grading used by institutions (percentage marking or shorter grade-point scale). This is an interesting example of the sort of practitioner led research that is needed to investigate some of the issues described in this article.

Quality of student work

Overwhelmingly the comments about student achievement are positive, with 50% commended. That the quality of student work comes in for so much praise speaks highly of the UK maths provision. Sometimes this evidence is supplemented by employers and former students. A substantial number of criticisms have been made however (20%). Disparity between the achievements of different groups of students have been noted, and some poor performance by relatively highly qualified students. Reviewers have cited cases of the level of work being over-stated, or substantial variability in the level of work demonstrated.

Use of assessment results in MIS

Good practice was noted where assessment results were used to track students' progress and support them, or identify areas of development needed in the curriculum. Weaknesses in this area were noted at a number of institutions. Normally complaints related to the actual difficulty of getting the assessment data, especially for particular sub-cohorts of students.

Conclusion

This article is only intended to flag up the main points concerning assessment arising from Subject Review. The wealth of good practice in student assessment that exists across the sector will be reported more fully in subsequent publications. However, there are also some issues remaining that would benefit from practitioner led educational research and development. The recently established FDTL Phase 4 Assessment Projects Support Group (<http://www.ncteam.ac.uk>) should provide a useful resource in supporting such projects.

One clear message that comes across from the work so far is that providers with good assessment records (in Subject Review) often rely on a cadre of experienced and dedicated teachers, many of whom may be retiring in the next five years [17]. In transferring the

responsibility for student assessment to those they leave behind there is an urgent need for effective staff and educational development in this area.

REFERENCES

1. Cox, W and Bidgood, P (2001) *Where are we now? – learning from Subject Review*, MSOR Connections **1**, 3, 15
2. *CAA mathematics series*, MSOR Connections, **1**, 4, 59
3. MSOR Connections (2002), **2**, 3, 4
4. Brown, G, Bull, J and Pendlebury, M (1997) *Assessing student learning in higher education*, Routledge, London
5. Heywood, J (2000) *Assessment in higher education*, Jessica Kingsley Publishers, London
6. Brown, G (2001) *Assessment: a guide for lecturers*, LTSN Generic Centre, York
7. Smith, G; Wood, L; Coupland, M; Stephenson, B; Crawford, K; Ball, G (1996) *Constructing mathematical examinations to assess a range of knowledge and skills* IJMEST **27**, 1, 65-77
8. Galbraith, P and Haines, C (2000) *Conceptual mis(understandings) of beginning undergraduates*, IJMEST **31**, 5, 651-678
9. Kahn, P and Kyle, J (Eds) (2002) *Effective learning and teaching in Mathematics and its applications*, Kogan Page, London
10. MSOR Connections August 2002, **2**, 3, 54
11. Yorke, M et al (2002) *Does grading method influence honours degree classification?* AEHE, **27**, 3, 269-279
12. Bishop, P (2002) *Report on Assessment for a purpose workshop*, MSOR Connections, **2**, 3, 54
13. Beevers, C and Bishop, P *Matching assessment to learning*, MSOR Connections, **1**, 3, 29
14. QAA (2002) Benchmark statement for MSOR
15. Bishop, P and Hibberd, S (2001) *A national approach to diagnostic testing*, MSOR Connections, **1**, 4, 56
16. Atkinson, R et al, in Undergraduate Mathematics Teaching Conference (UMTC) 2001, edited by P Egerton, page 49
17. Middleton, W, in UMTC 2001, edited by P Egerton, page 12