

The following paper by Ken Houston was published in *The Teaching and Learning of Mathematics at University Level: An ICMI Study*, (Derek Holton, Editor), Kluwer Academic Publishers: The Netherlands, 2001, pages 407-422. (ICMI is the International Commission on Mathematics Instruction.)

The paper is posted here with kind permission of Kluwer Academic Publishers, editor Derek Holton, and author Ken Houston. The volume containing the paper is volume 7 of the New ICMI Study Series, co-editors Bernard R. Hodgson (Canada) and Hyman Bass (USA).

Derek Holton (Ed.), *The Teaching and Learning of Mathematics at University Level: An ICMI Study*, 407-422.
© 2003 Kluwer Academic Publishers. Printed in the Netherlands.

KEN HOUSTON

ASSESSING UNDERGRADUATE MATHEMATICS STUDENTS

1. INTRODUCTION

Any discussion of assessment must necessarily include a discussion of the curriculum, how it is designed and organised, and what it contains. It must examine the aims of the course that students are taking, and the objectives set for that course and the individual modules that comprise the course. (Here I am using terminology common in the UK. The 'course' students take is 'the whole thing', the 'programme'. A course in this sense consists of 'modules' or 'units', commonly called 'courses' in the USA, so beware of confusion!) The discussion must consider who is doing the assessing, why they are doing it, what they are doing and how it is being done. It must consider how assessors become assessors and how those assessed are prepared for assessment. And it must consider if the assessment is valid and consistent, and if it is seen to be so.

It might also be useful at this stage to define what we mean by a 'mathematician'. There is a real sense in which almost everyone could be described as a mathematician in that they make use of some aspect of mathematics – be it only arithmetic or other things learnt at primary/elementary school. The term could be used of those who have taken a first degree in mathematics and who use it in their employment. Or it could be reserved only for those who have a PhD and who are doing research in pure mathematics or an application of mathematics. We will use the middle of the road term. In other words, a mathematician will be one who has studied the subject at least to bachelors degree standard (and of course that varies across the world!), and who is using some aspect of advanced mathematics in their work. Such people could join a professional or learned society such as the UK based Institute of Mathematics and its Applications. So we are primarily concerned with the higher education of these people who can rightly be considered to be professional mathematicians. But also there are many disciplines wherein mathematics is an extensive and substantial component of study. Examples are physics or electronic engineering. The mathematical education of professionals in such fields as these could also come under the remit of this article in that many of the suggestions made could enhance the teaching, learning and assessment of students in these fields.

Traditionally assessment in higher education was solely summative and consisted of one or more time-constrained, unseen, written examination papers per module. A typical, and in some places predominant, purpose of assessment was to put students in what was believed to be rank order of ability. Students were, perhaps,

407

Derek Holton (Ed.), *The Teaching and Learning of Mathematics at University Level: An ICMI Study*, 407—422.

© 2003 Kluwer Academic Publishers. Printed in the Netherlands.

asked to prove a theorem or to apply a result, or to see if they could solve some previously unseen problem. Generally this method succeeded in putting students in a rank order and in labelling them excellent, above average, below average or fail. But was it rank order of ability in mathematics or rank order of ability to perform well in time-constrained, unseen, written examination papers? Sadly it was the latter, and while the two may coincide, this is not guaranteed. Taking time-constrained, unseen, written examination papers is a rite of passage, which students will never have to do again after graduation and which bears little relationship to the ways in which mathematicians work. While it is true that working mathematicians are sometimes under pressure to produce results to a deadline, the whole concept of time-constrained, unseen, written examinations is somewhat artificial and unrelated to working life.

It is in this context that people started to think about change, change in the way courses are designed and organised, change in the way course and module objectives are specified and change in the way students are assessed and in the way the outcomes of assessment are reported. It is usually the case that ‘what you assess is what you get’, that is, the assessment instruments used determine the nature of the teaching and the nature of the learning. Learning mathematics for the principal purpose of passing examinations often leads to surface learning, to memory learning alone, to learning that can only see small parts and not the whole of a subject, to learning wherein many of the skills and much of the knowledge required to be a working mathematician are overlooked. In time-constrained, unseen, written examinations no problem can be set that takes longer to solve than the time available for the examination. There are no opportunities for discussion, for research, for reflection or for using computer technology. Since these are important aspects of the working mathematician’s life, it seems a pity to ignore them. And it seems a pity to leave out the possibilities for deep learning of the subject, that is, learning which is consolidated, learning which will be retained because it connects with previous learning, learning which develops curiosity and a thirst for more, learning which is demonstrably useful in working life.

This is, of course, a caricature of ‘traditional’ assessment, but it is not too far from the truth, and it brings out the reasons why some people in some societies became unhappy with university and college education. Consequently those who educate students now pay attention to stating aims and objectives, to redesigning curricula and structures and to devising assessment methods which promote the learning we want to happen and which measure the extent to which it has happened. And they pay attention to the need to convince students and funding bodies that they are getting good value for their investment of time and money.

The discussion on course design and assessment is also tied up with the discussion on ‘graduateness’. What is it that characterises college or university graduates and distinguishes them from those who are not? Is it just superior knowledge of a particular topic, or is it more than that? It is, of course, more than that. It is not easy to define or even to describe, but it has to do with an outlook on life, a way of dealing with problems and situations, and a way of interacting with other people. (This is not to denigrate the learning that non-college graduates get from ‘the university of life’, nor to suggest that they are inferior as people. It is to do

with considering the ‘added value’ of college or university education.) Traditionally graduateness was absorbed, simply through the university experience, but now that we have systems of mass education in many countries of the world, we need to pay attention to the development of graduate attributes in students so that they do, indeed, get value for money. In many instances, and mathematics is no exception, it is the ‘more than’ that is important when it comes to finding and keeping employment. Subject knowledge is important but so also are personal attributes. It is highly desirable that students develop what have come to be known as ‘key skills’ while they are undergraduates, and not just because employers are saying that the graduates they employ are weak in this area. Innovative mathematics curricula seek to do this by embedding the development of key skills in their teaching and learning structures. (Key skills are often described as employability skills or transferable skills. They include such skills as written, oral and visual communication, time management, group-work and team-work, critical reflection and self assessment, and computer and IT, and aural skills.)

Who are the stakeholders in an undergraduate’s education? First and foremost are the students themselves. They are investing time and effort and they want to know that they are getting a return on this investment. Most of them realise that it is not enough for them to be given a grade; they know that they have to earn it. So they need to know what performance standards are required and they need to be able to recognise within themselves whether they have achieved these standards or not. This raises the question of self-assessment and ways of promoting self-assessment. Giving ‘grades that count’ is one way of encouraging students to carry out tasks.

The next stakeholder to consider are the teachers. It is their job to enable learning and so they need to know what learning has taken place. Financial sponsors of students are also stakeholders. They, too, want to know if they are getting a good return on their investment. Finally, in the stakeholder debate, there is a demand from society, students themselves, universities, prospective employers, that students be summatively assessed, ranked and labelled in such a way that they may be measured, not just against what they are supposed to have learned, but also against their peers across the world.

This chapter will consider all of these features, but will focus on assessment, as that is its theme. It will look at the purposes and principles of assessment and then it will move on to consider the aims and objectives of courses and modules. Innovative methods of assessment will be reviewed and discussed, and this will be the biggest part of the chapter. Ways of disseminating information about new assessment practices will be discussed, as will obstacles to change. Finally pertinent research issues will be mentioned. The chapter will close with an annotated bibliography of pertinent books and papers dealing with these issues.

2. PRINCIPLES AND PURPOSES OF ASSESSMENT

Perhaps the only principle that should be applied is ‘fitness for purpose’. To achieve this, assessment methods should be intimately related to the Aims and Objectives of the Module under consideration. And it should be born in mind that

the assessment methods used will influence the learning behaviour of students to a considerable extent.

There are a number of purposes of assessment that should be considered:

1. to inform learners about their own learning.
2. to inform teachers of the strengths and weaknesses of the learners and of themselves so that appropriate teaching strategies can be adopted.
3. to inform other stakeholders – society, funders, employers including the next educational stage.
4. to encourage learners to take a critical-reflective approach to everything that they do, that is, to self assess before submitting.
5. to provide a summative evaluation of achievement.

3. AIMS AND OBJECTIVES

Aims and objectives should be established both for a course and for each of the modules that comprise the course. The aims of a course are statements that identify the broad educational purposes of the course and may refer to the ways in which it addresses the needs of the stakeholders. Here are some examples; there are, of course, many more and each provider must write their own:

1. To provide a broad education in mathematics, statistics and computing for students who have demonstrated that they have the ability or who are considered to have the potential to benefit from the course.
2. To develop knowledge, understanding and experience of the theory, practice and application of selected areas of mathematics, statistics, operations research and computing so that graduates are able to use the skills and techniques of these areas to solve problems arising in industry, commerce and the public sector.
3. To develop key skills.
4. To provide students with an intellectual challenge and the practical skills to respond appropriately to further developments and situations in their careers.
5. To prepare students for the possibility of further study at post graduate level, including a PhD programme or a teacher training programme.

It would be necessary to indicate how each of the modules selected for a course helps to achieve the aims of the course. The aims of the individual modules should ‘map’ to the overall aims of the course. Objectives are statements of the intended learning outcomes that would demonstrate successful completion of the course or module, and that would warrant progression through the course and the eventual award of a degree. Module objectives should identify the knowledge, skills and attributes developed by a module, and course objectives should identify the knowledge, skills and attributes developed by the totality of modules selected for the course. Objectives may include reference to subject knowledge and understanding, cognitive skills, practical skills and key skills. They should be clearly relevant to fulfilling the aims and, above all, they should be assessable, that is, we should be

able to devise assessment instruments that allow students to demonstrate that they have achieved the learning intended, and, if appropriate, to what extent. Here are some examples of course objectives: -

On completion of their studies graduates will have:

1. an understanding of the principles, techniques and applications of selected areas of mathematics, statistics, operations research and computing.
2. the ability and confidence to analyse and solve problems both of a routine and of a less obvious nature.
3. the ability and confidence to construct and use mathematical models of systems or situations, employing a creative and critical approach.
4. effective communication skills using a variety of media.
5. effective teamwork skills.

A course document should demonstrate how the aims and objectives of the constituent modules contribute to the overall course aims and objectives. Here is an example of the aims and objectives of a module, taken from an introductory module on mathematical modelling. (These aims and objectives are those of module MAT112J2, University of Ulster. Full details may be read under 'Syllabus Outline' at <http://www.infj.ulst.ac.uk/~cdmx23/mat112j2.html>.) Note that an indication of the method of assessment of each objective is given.

Aims: The aims of this module are to:

1. enable students to understand the modelling process, to formulate appropriate mathematical models and to appreciate their limitations.
2. develop an understanding of mathematical methods and their role in modelling.
3. study a number of mathematical models.
4. develop in students a range of key skills.

It can be seen how these module aims help to meet the aims of the course listed above. Thus this module contributes to developing mathematical understanding, problem solving, and key skills.

Objectives: On completion of this module, students should be able to:

1. Formulate mathematical models and use them to solve problems of an appropriate level. (Assessed by coursework and written examination.)
2. Solve simple differential equations using calculus and computer algebra systems. (Assessed by written examination.)
3. Describe and criticise some mathematical models. (Assessed by coursework.)
4. Work in groups and report their work in a variety of media. (Assessed by coursework.)
5. Work both independently and in support of one another. (Assessed by coursework.)
6. Demonstrate other key skills. (Assessed by coursework.)

Again, it can be seen how these module objectives map to the course objectives listed above. There are references to the assessment of mathematical techniques, the construction and use of mathematical models, and key skills.

Of course, aims and objectives are not created in a vacuum. They evolve from the previous and present experiences of the lecturing staff who design the course and its constituent modules, and they are reviewed and modified from time to time as circumstances permit or demand. Nevertheless, the objectives for each module and for the course as a whole should be stated and essentially should be a form of contract between the lecturer and the students. Furthermore, detailed assessment criteria should be drawn up so that lectures have a well-defined framework in which to work and students have clear guidelines to what they have to do in order to succeed. This contractual arrangement does to some extent limit the power traditionally wielded by lecturers. This is a necessary and desirable consequence of the innovations described in this paper. It makes students more powerful in the right context, namely their own learning, in that it does require students to take more responsibility for themselves.

4. EXTERNAL ASSESSMENT OR EVALUATION

While this paper is primarily about the assessment of student learning, it may be appropriate to mention current developments in the evaluation of institutions, their courses and modules, and the teaching and other staff who deliver these. In the UK, for example, the government agency, The Quality Assurance Agency for Higher Education has a remit to review the quality of provision of education by institutions. Mathematics courses were reviewed between 1998 and 2000, along with several other subjects. The whole of university life is covered in approximately six-year cycles. Institutions are required to write an evidence-based self-assessment document (the SAD), which a visiting team of reviewers will scrutinise and make a judgement on. The SAD outlines the aims and objectives of the provision and provides details of the physical and human resources available. It then gives the institution's own assessment of its quality of provision under six headings:

1. Curriculum Design, Content and Organisation;
2. Teaching, Learning and Assessment;
3. Student Progression and Achievement;
4. Student Support and Guidance;
5. Learning Resources;
6. Quality Management and Enhancement.

Evidence to support the claims made in the SAD must be provided and may be found in documents and in observation of teaching.

This peer-review process (the reviewers are selected from other, similar institutions) devours a considerable amount of academic time and energy, and it remains to be seen if the improvements justify the cost. It is part of the general move in society to satisfy the public demand for public accountability of public funds. On

the positive side, it has encouraged institutions to think about their course provision in a way that would be new to many.

5. METHODS OF ASSESSMENT

Once the learning outcomes or objectives have been articulated, suitable assessment methods have to be selected. (In practice, the articulation of objectives and the selection of assessment methods will proceed hand-in-hand.) This should be done in such a way as to ensure that the assessment methods are appropriate and allow students to demonstrate positive achievement. There should be transparent assessment criteria, which should be explained to students, if possible with examples of good work and not so good work harvested from previous cohorts, or descriptions of excellent, median and pass level performance. Ideally the assessment criteria should be drawn up in debate with the students, without sacrificing the lecturer's expertise and experience. Assessment should blend with the teaching and learning pattern. This section will now review some assessment practices that have been developed and used successfully.

5.1 *Individual project work*

Project work, both individual project work and group project work, is used widely. This has been a feature of many undergraduate mathematics courses for over twenty-five years, so it can hardly be described as innovative. Individual projects are often given to final year students and are substantial pieces of work. The very least the project is worth is about one sixth of final year studies, but it may be worth more than that. The topics set for investigation can be quite demanding and give scope for considerable initiative and independent work by students. Projects demand research and investigation and the production of a written report and some sort of presentation, such as a seminar, a poster or a *viva voce* examination. Students learn to conduct research and to organise information and present it cogently.

But now comes the hard part – assessing it reliably and validly – and ensuring that students know how it is to be assessed. If students know the assessment criteria and have some idea of what constitutes good or not so good work, then they are in a better position to assess their own work before submitting it, and in a better position to assess one another's work in a peer support activity. Project work like this is a good method for assessing many of the objectives outlined above that lead to the development of 'the way of life' of a working mathematician in whatever guise they may find themselves.

Experienced project assessors can usually come to an accurate judgement of a student's work fairly quickly and can defend that judgement to their peers. But there still is an element of subjectivity in this and, to remove as much of this as possible and to achieve consistent marking by several assessors, consultation and training are necessary. The team of assessors should develop assessment criteria, should trial their use and should analyse and reflect on their judgements. In this way hard markers and lenient markers will be identified, and all will learn how to apply the

detail of the criteria. Inexperienced assessors need this sort of training exercise at the start of their careers.

5.2 Group project work

Group project work is often introduced at an earlier stage in a student's career. Again this gives opportunity for encouraging research, investigation and communication. But it also introduces students to group work and the problems associated with that. Often the internal working of the group can best be assessed by the members of the group themselves. This can be accomplished through confidential self and peer assessment. Sometimes it is more appropriate for the lecturer to observe the working process; this method has the added advantage that the lecturer can intervene in crisis situations. Assessors face dilemmas when assessing project work carried out by groups. If the same grade is given to each member of the group, then some may benefit and some may suffer from the work, or lack of it, of their peers, and this could be considered to be unfair, both by students and by society. Experience shows that this is a price worth paying. The dilemmas can be overcome by including an element of confidential, within-group, peer assessment, by observing group work, and by ensuring that students experience a good mix of group and individual assessment methods throughout their course. In working life, after all, group leaders often carry the blame for the poor performance of their group. By that stage, of course, they will be much more experienced and will have more control over their staff, but it is a useful lesson for students to discover the difficulties of working with other people, provided it is not disastrous to the overall outcome of their time at university.

5.3 Variations on written examinations

Variations on the theme of written examinations have been tried. These include open book examinations, seen examinations wherein students are given the questions some time in advance and they prepare their answers to them, examinations conducted in a computer laboratory with ready access to computer algebra systems and other mathematical software, and examinations which involve conceptual questions.

5.4 Comprehension tests

Some experiments on the use of Comprehension Tests have been carried out. This method of assessment is widely used in other subjects. Students are given an article or part of a book to read in advance. They study it very carefully and then take an unseen written paper, which is designed to explore the extent to which each student has comprehended the article. This can be useful for assessing a student's understanding of mathematical processes. Furthermore it encourages students to read critically and reflectively, to try to get into the mind of the author, and to think

deeply about the topic of the article. It helps them to see that mathematics is alive and active in some contemporary context.

5.5 *Journal writing*

Student journal writing through the course can be used to help diagnose learning difficulties and to address these at an early stage. Students may be given time at the end of a teaching session to reflect on their learning during that session and to write down their thoughts and feelings, their worries and concerns, what they have learnt and what they are having difficulty with. Or they may be asked to do this overnight, thus giving them at least a little time to digest the day's work. The journals should be read frequently by the lecturer so that formative feedback can be given in good time and appropriate intervention strategies introduced if necessary.

Other strategies have been developed and used, such as brief, 'one-minute' quizzes or student-written summaries of key points learned (or not) at the end of class periods to provide feedback to instructors, particularly at early stages of modules. Student portfolios, student lectures and combined written-oral examinations are other strategies that have been used to good effect

6. DISSEMINATION OF INNOVATIVE IDEAS AND CONCLUSIONS OF RESEARCH STUDIES

The impetus for change and innovation usually comes from individuals who are dissatisfied with what they have been doing. They will have experimented, preferably with the approval of their head of department but sometimes covertly, and evaluated the effects of their ideas, and then adopted or scrapped them. The wider mathematics community can be informed about these developments in the same ways that research findings are disseminated, that is, by word of mouth at seminars and conferences, and by publication. It is helpful if papers relating to teaching and learning research are included in mainstream mathematics conferences and journals. Then lecturers who would never dream of attending a 'teaching' conference or reading a 'teaching' journal might just be exposed to these ideas and might be persuaded to accept them and adopt them.

Very occasionally a charitable foundation or a government agency will fund the production and dissemination of material relating to teaching innovation. Some examples are given in the annotated bibliography at the end of this chapter.

7. OBSTACLES TO CHANGE AND STRATEGIES FOR OVERCOMING THEM

Ignorance and prejudice are, perhaps, the greatest obstacles to change. Lack of resources is another. Many teachers in higher education will not have attended a course on teaching as part of their pre-service or in-service preparation for the job. (Most will have completed a PhD or equivalent and will be well versed in research

methodology.) And so they will most probably teach as they themselves were taught. They are ignorant (in the nicest possible sense of the word!) of new ideas and new scholarship in student learning. Overcoming ignorance is relatively easy but requires an extensive programme of dissemination of ideas targeted particularly at new lecturers. One strategy being introduced in the UK is the requirement of many institutions for new lecturing staff to complete a post graduate certificate in university teaching. Usually this will be a two-year, part-time course delivered by the lecturer's own institution (or local consortium) and completing it successfully is a requirement of probation. Courses will usually include modules on generic and subject specific teaching and learning and the assessment instruments will include a portfolio of work relating to the lecturer's own teaching. Another strategy introduced in the UK is the recommendation that all lecturers join, and maintain their membership of, the newly constituted Institute for Learning and Teaching in Higher Education (ILT). Membership of the ILT gives public recognition that the member has had training as a teacher and continues to develop professionally. It will function as a professional association. (For two years there will be a special route to membership for experienced teachers, who will not be required to undergo an initial training programme but will, instead, submit a relatively short document outlining their experience as teachers and including a section of critical reflection on their work. See Mason, this volume, pp. 529-538.) It helps greatly if the head of department and other senior officers in the institution are sympathetic to the aims of such a programme, are knowledgeable about developments in teaching and learning, and support and encourage their colleagues to overcome their ignorance.

Prejudice is much harder to overcome. Prejudice is when a person, in full knowledge of developments, still rejects them unreasonably and out-of-hand, just because they are innovations or perhaps because once they encountered badly written arguments for change or are suspicious of research in education. This requires greater evangelistic effort. Research findings must be carefully presented and arguments for innovation persuasively written. Personal contacts, one-to-one over a meal or a drink, are good opportunities for this. Again, encouragement from heads of department and higher is very valuable.

Resource issues are important also. Recent years have seen resources diminish in universities all over the world. Classes are bigger and lecturers have conflicting and very strong demands on their time, particularly the demand to carry out high quality research. Money to pay for professional development is in short supply. There are no easy answers to the problem of lack of resources. It is a matter of commitment and priority, particularly on the part of heads of department and other resource managers in institutions. If the leader of a unit is committed to innovation and development in teaching and learning, then it is more likely to happen. One possible help might come from the bodies that determine the resources given to universities or departments to conduct research. If they were to allow research into the teaching and learning of a subject to have equal status with research into the subject itself, and if they were to allocate funds for this, then it is likely that more attention would be paid to teaching development. Also other carrots within institutions, such as promotion criteria which include good teaching, would help to stimulate the developments outlined in this paper.

Perhaps one of the more serious obstacles to be faced, as regards change and innovation in assessment at the university level, is that there may be genuine conflicts of interest between different stakeholders and parties. For instance, there may be a clash between, on the one hand, academics who tend to insist on the (in)formative purpose of assessment and on the ensuing necessity of multi-faceted and complex assessment instruments for validly capturing a fair range of knowledge and skills with their students, and, on the other hand, institution heads and administrators who tend to insist on the summative or ranking aspects of assessment in order for the institution to live up to common expectations in the environment or to make a positive appearance in a highly competitive 'university market'. However, even if this sort of clash is not present, clashes may arise if, as is often the case, the innovative assessment methods advocated or adopted by academics turn out to be considerably more time consuming or resource intensive than the traditional methods. At times when university funding is scarce heads and administrators may be inclined to counteract the use of such methods, not because of scepticism towards their relevance but simply because of the resources they consume. Such sorts of clashes are of an objective nature and cannot always be easily reconciled.

8. PERTINENT RESEARCH ISSUES

As mentioned above, there are sceptics in universities who are not yet convinced that teaching innovations are necessarily a good thing, but who may still be receptive to persuasive arguments and research findings. So research which seeks to evaluate innovative teaching methods and which demonstrates that aims and objectives are being met, is needed. Of course, the teaching developments themselves must have a rationale which is based on research into student learning. Another field of study is the robustness of the assessment methods themselves.

There are a number of internationally renowned teams who have published widely their research into student learning. Most of the work on assessment has been carried out by the Assessment Research Group (ARG) in the UK and some members of the International Community of Teachers of Mathematical Modelling and Applications (ICTMA). The work of the ARG is reported at length elsewhere in this volume (see Haines and Houston, this volume, pp. 431-442). Their main work has been to develop, test and evaluate robust methods for assessing several different forms of student project work and associated communication skills. They were also the nucleus of a group who received UK government funding to develop and disseminate resource material relating to innovative learning and assessment. Since some of this work related to teaching mathematical modelling, members of ARG are also active in the ICTMA, and some original research is published in the ICTMA series of books and conference abstracts.

But every teacher can be a researcher in their own classroom, picking up good ideas, developing them, evaluating them and then telling the world about them. This is actually quite an exciting thing to do and people who do it get a buzz from the experience which invigorates themselves, their teaching and their students.

ANNOTATED BIBLIOGRAPHY

- American Association of University Professors (1990). Mandated Assessment of Education Outcomes. *Academe*, Nov./Dec. Discusses impact of mandated assessment on traditional arenas of professorial autonomy; focuses on five assessment issues (institutional diversity, skills, majors, value-added, and self-improvement). Concludes with recommendations for learning to live with mandated assessment.
- Angelo, Thomas A. and Cross, K. Patricia (1993). *Classroom Assessment Techniques: A Handbook for College Teachers*, 2nd Ed, San Francisco: Jossey-Bass Publishers. This text focuses on formative classroom assessment. In addition to describing what classroom assessment is and how one might plan and implement classroom assessment tasks, the authors present 50 different classroom assessment techniques, many of which can be used in or modified for the mathematics classroom.
- Astin, Alexander, Banta, Trudy W., Cross, Patricia K., et al. (1992). *Principles of Good Practice for Assessing Student Learning*. Washington, DC: American Association of Higher Education. Nine principles for assessing student learning developed by the long-standing annual Assessment Forum of the American Association of Higher Education (AAHE).
- Ball, G, Stephenson, B, Smith, G, Wood, L, Coupland, M, and Crawford, K. (1998). Creating a Diversity of Mathematical Experiences for Tertiary Students. *Int J Math Edu Sci Technol*, 29, 827-841.
- Bass, Hyman (1993). Let's Measure What's Worth Measuring. *Education Week*, October 27, 32. An 'op-ed' (opinion) column supporting Measuring What Counts from the Mathematical Sciences Education Board (MSEB). Stresses that assessments should (a) reflect the mathematics that is most important for students to learn; (b) support good instructional practice and enhance mathematics learning; and (c) support every student's opportunity to learn important mathematics.
- Benjamin, Ernst (1990). The Movement to Assess Students' Learning will Institutionalize Mediocrity in Colleges. *Chronicle of Higher Education*, July 5. A brief op-ed column criticizing the indefensible consequences for higher education of rapidly spreading accountability systems that rely on narrow tests.
- Benjamin, Ernst (1994). From Accreditation to Regulation: The Decline of Academic Autonomy in Higher Education. *Academe*, July/Aug., 34-36. A worried analysis by the retired general secretary of the American Association of University Professors (AAUP) concerning the impact of increased regulation based on accountability measures that, Benjamin believes, might distort traditional goals of the academy.
- Berry, J. and Haines, C.R. (1991). *Criteria and Assessment Procedures for Projects in Mathematics*. Plymouth: University of Plymouth. The first of four reports written by the UK Assessment Group (ARG). It is a report of a workshop held in 1991, which aimed to review the assessment schemes and assessment criteria currently in use in UK universities and to begin to develop robust criteria-based assessment procedures for a wide range of topics. It describes the use of the FACETS data analysis package to consider the differing effects of the criteria themselves and the ways in which the assessors used them.
- Berry, J. and Houston, K. (1995). Students using Posters as a means of Communication and Assessment. *Educational Studies in Mathematics*, 29, 21-27. Gives a thorough literature review of the use of posters by students in higher education generally, and suggests ways in which student posters could be used and could be assessed in mathematics classes.
- Bok, Derek (1986). Toward Higher Learning: The Importance of Assessing Outcomes. *Change*, Nov./Dec., 18-27. A classic short essay by then Harvard President Derek Bok outlining the benefits to higher education of assessing the accomplishments and value of a college education.
- Burton, L. and Izard, J. (1995). *Using FACETS to Investigate Innovative Mathematical Assessments*. Birmingham: University of Birmingham. The last of four reports written by ARG. It reports on the 1994 workshop and discusses continuing issues in assessing student project work. It deals with reliability of marking schemes, self, peer and tutor marked assignments, equity, and criteria for the assessment of students' posters.
- Committee on the Undergraduate Program in Mathematics (CUPM) (1995). Assessment of Student Learning for Improving the Undergraduate Major in Mathematics. *Focus: The Newsletter of the Mathematical Association of America*, 15:3 (June), 24-28. Recommendations from the Mathematical Association of America (MAA) for departments of mathematics to develop a regular 'assessment cycle' in which they (1) set student goals and associated departmental objectives; (2) design instructional strategies to accomplish these objectives; (3) select aspects of learning and related assessments in which quality will be judged; (4) gather assessment data, summarize this information,

and interpret results; and (5) make changes in goals, objectives, or strategies to ensure continual improvement.

- Edgerton, Russell (1990). Assessment at Half Time. *Change*, Sept./Oct., 4-5. A brief summary of the political landscape of assessment in higher education by the president of the American Association of Higher Education (AAHE). Claims that state pressures for accountability will continue, but that if institutions define assessment in worthy terms the faculty will find the effort worthwhile.
- Ewell, Peter T. (1997). Strengthening Assessment for Academic Quality Improvement. In Marvin W. Peterson, David D. Dill, and Lisa A. Mets (Eds.), *Planning and Management for a Changing Environment: A Handbook on Redesigning Postsecondary Institutions*, pp. 360-381. San Francisco: Jossey-Bass Publishers. Historical survey of assessment efforts in the U.S. during the last decade in the context of increased accountability requirements, decreased financial resources, and increased experience with assessment on college campuses. Discusses relation of assessment to academic planning and syndromes to avoid.
- Ewell, Peter T. and Jones, Dennis P. (1996). *Indicators of Good Practice in Undergraduate Education: A Handbook for Development and Implementation*. Boulder, Co: National Center for Higher Education Management Systems (NCHEMS). Intended to provide colleges and universities with guidance in establishing an appropriate system of indicators of the effectiveness of undergraduate instruction, and to build on this foundation by cataloguing a range of exemplary indicators of 'good practice' that have proven useful across many collegiate settings.
- Ferren, Ann (1993). *Faculty Resistance to Assessment: A Matter of Priorities and Perceptions*. Commissioned paper prepared for American Association of Higher Education. Analyzes faculty priorities to help understand why assessment is rarely valued by faculty. Argues that assessment must derive from widely agreed goals, must be connected to clear outcomes that the faculty see as beneficial, and must not be simply added to already overburdened faculty loads.
- Frechtling, Joy A., 1995. *Footprints: Strategies for Non-Traditional Program Evaluation*. Washington, DC: National Science Foundation. A series of papers suggesting diverse strategies for assessing the impact of funded programmes both short- and long-term, both intended and unintended.
- Gaither, Gerald H. (1995). *Assessing Performance in an Age of Accountability: Case Studies*. San Francisco: Jossey-Bass Publishers. Case studies from several states and public institutions about the shift from campus-based assessment in the 1980s to state-based accountability systems in the 1990s.
- Glassick, Charles E., Huber, Mary T. and Maeroff, Gene I. (1997). *Scholarship Assessed: Evaluation of the Professoriate*. Carnegie Foundation for the Advancement of Teaching. San Francisco: Jossey-Bass Publishers. Companion to Ernest Boyer's widely-cited *Scholarship Reconsidered*: Priorities of the Professoriate, this report outlines standards for evaluating scholarship that transcend differences among disciplines: clear goals, adequate preparation, appropriate methods, significant results, effective presentation, and reflective critique.
- Gold, Bonnie, Keith, Sandra and Marion, William (Eds.), (1999). *Assessment Practices in Undergraduate Mathematics*. Washington, DC: Mathematical Association of America. A collection of over fifty brief reports from dozens of different U.S. colleges and universities providing a wide variety of methods of assessing the major, teaching, classroom practice, the department's role, and calculus reform.
- Haines, C.R. and Dunthorne, S. (Eds.) (1996). *Mathematics Teaching and Learning-Sharing Innovative Practices*. London: Arnold. This resource pack is a collection of articles describing innovative practices in teaching and assessment. It was written by mathematics lecturers from a consortium of UK universities, including members of ARG.
- Haines, C.R. and Izard, J. (1995). Assessment in Context for Mathematical Modelling. In C. Sloyer, W. Blum and I. Huntley (Eds.). *Advances and Perspectives in the Teaching of Mathematical Modelling and Applications* pp. 131-149. Yorklyn, Delaware: Water Street Mathematics. Credible assessment schemes measure evidence of student achievement, as individuals or within a group, over a wide range of activities. This paper shows that item response modelling can be used to development rating scales for mathematical modelling. It draws on the work of ARG and work in Australia.
- Haines, C.R., Izard, J. and Berry, J. (1993). *Awarding Student Achievement in Mathematics Projects*. London: City University. The second of four reports written by ARG. It investigates in depth the use of assessment criteria for judging oral presentations by students of their project work. It also proposes criteria for the assessment of written reports of different types of student project work.

- Hilton, Peter (1993). The Tyranny of Tests. *American Mathematical Monthly*, April, 365-369. Several suggestions for 'reducing the distorting effect' which tests exert, principally on undergraduate mathematics.
- Houston, S.K. (1993). *Developments in Curriculum and Assessment in Mathematics*. University of Ulster. This pamphlet contains papers presented at a one day symposium at the University of Ulster following the 1993 meeting of ARG.
- Houston, S.K. (1993). Comprehension Tests in Mathematics (I and II), *Teaching Mathematics and its Applications*, 12, 60-73 and 113-120. These are the original papers describing the use of comprehension tests in mathematics.
- Houston, S.K. (1995). Assessing Mathematical Comprehension. In C. Sloyer, W. Blum, and I. Huntley (Eds.), *Advances and Perspectives in the Teaching of Mathematical Modelling and Applications*, pp. 151-162. Yorklyn, Delaware: Water Street Mathematics. This paper examines the rationale for comprehension tests in mathematics, and outlines possible aims and objectives. It describes the author's experiences in setting and using such tests and outlines the extent of their use in secondary and tertiary education in the UK and extends the work reported in Houston (1993).
- Houston, S.K. (1997). Evaluating Rating Scales for the Assessment of Posters. In S.K. Houston, W. Blum, I. Huntley and N.T. Neill, (Eds.), *Teaching and Learning Mathematical Modelling*, pp. 135-148. Chichester: Albion Publishing (now Horwood Publishing). Deals with the use of posters by university students as a means of communication and as a vehicle for assessment. There is a summary of a literature review and a rationale for the activity. The author claims that it is an enjoyable activity, which is beneficial for students. The main purpose of the paper is to describe the development and evaluation of assessment criteria and rating scales.
- Houston, S.K., Haines, C.R. and Kitchen, A. (1994). *Developing Rating Scales for Undergraduate Projects*. University of Ulster. The third of four reports written by ARG. It reports on the 1993 workshop, giving details of assessment criteria (or descriptors) for the assessment of written reports on projects in pure mathematics, mathematical modelling, statistical investigations and investigations of a more general nature. The report describes how the group members trialled the criteria and how the data analysis led to the development of robust assessment procedures. It also introduces the use of criteria for the assessment of student posters.
- Izard, J. (1997). Assessment of Complex Behaviour as Expected in Mathematics Projects and Investigations. In S.K. Houston, W. Blum, I. Huntley and N.T. Neill, (Eds.), *Teaching and Learning Mathematical Modelling*, pp. 109-124. Chichester: Albion Publishing (now Horwood Publishing). No single assessment method is capable of providing evidence about the full range of achievement. This paper reviews the problems faced in devising better assessments to monitor learning, and provides practical suggestions for meeting these problems. The methods presented are applicable to traditional examinations, project and investigation reports, presentations and posters, judgements of performance and constructed projects, observations of participation, collaborative group work and ingenuity. The paper concludes with advice on monitoring the quality of the assessment process.
- Joint Policy Board for Mathematics (1994). *Recognition and Rewards in the Mathematical Sciences*. Providence, RI: American Mathematical Society. Discussion of faculty expectations in relation to institutional rewards. Findings include a general dissatisfaction with current methods of evaluating teaching as well as uncertainty about the weight of effective teaching in college expectations and rewards.
- Katz, Stanley N. (1994). Defining Education Quality and Accountability. *Chronicle of Higher Education*, November 16, A56. An op-ed statement by the president of the American Council of Learned Societies (ACLS). Urges that colleges and universities heed the wake-up call of assessment from elementary and secondary schools and figure out how to define educational quality in terms that are worthy of higher education.
- Linn, Robert L. and Herman, Joan L. (1997). *A Policymaker's Guide to Standards-Led Assessment*. Denver, CO: Education Commission of the States. Analysis of policy implications involved in shifting from norm-referenced assessments (which compare each students' performance to that of others) to standards-led assessments which incorporate pre-established performance goals, many of which are based on real-world rather than 'artificial' exercises.
- Madison, Bernard (1992). Assessment of Undergraduate Mathematics. In Lynn A. Steen (Ed.), *Heeding the Call for Change: Suggestions for Curricular Action*, Washington, pp. 137-149. DC: Mathematical Association of America. Analysis of issues, benefits, worries, and pressures associated

- with the increasing demand for assessment of undergraduate mathematics. A background paper preceding release of the CUPM report on assessment.
- Mathematical Sciences Education Board (1993). *Measuring What Counts: A Conceptual Guide for Mathematics Assessment*. Washington, DC: National Research Council. Intended primarily as advice for K-12 mathematics assessment, this report stresses the need for assessment to measure good mathematics, to enhance learning, and to promote access for all students to high quality mathematics.
- National Council of Teachers of Mathematics (1995). *Assessment Standards for School Mathematics*. Reston, VA: National Council of Teachers of Mathematics. This third and final volume in NCTM's original set of standards for school mathematics focuses on six standards: effective assessment should reflect appropriate mathematics, enhance learning, promote equity, be based on an open process, promote valid inferences, and fit together coherently.
- Niss, Mogens (Ed.) (1993). *Investigations into Assessment in Mathematics Education - An ICMI Study*. Dordrecht: Kluwer Academic Publishers. This book is one of two resulting from the ICMI Assessment Study. The book offers a variety of approaches to the conceptual, philosophical, historical, societal, and pedagogical investigation of assessment in mathematics education, by prominent mathematics educators from Europe, North America and Australia. Both survey chapters and specific empirical or theoretical studies are included in the book.
- Open University Course Team (1998). *Assessment of Key Skills in the Open University Entrance Suite*. MU120, MST121, MS221. Open University, Milton Keynes.
- Romer, Roy (1995). *Making Quality Count in Undergraduate Education*. Denver, CO: Education Commission of the States. Report by the then-Governor of Colorado on behalf of all U.S. state governors concerning what parents and students expect of higher education and what research says about the characteristics of high-quality undergraduate education. Concludes with recommendations for steps to make higher education more accountable to its public purposes.
- Schilling, Karen Maitland and Schilling, Karl L. (1993). Professors Must Respond to Calls for Accountability. *Chronicle of Higher Education*, March 24, A40. An op-ed column arguing that faculty must take seriously the public's demand for evidence that students are learning, and learning the right things. Suggests portfolio assessment as an effective strategy.
- Schoenfeld, Alan (1997). *Student Assessment in Calculus*. Washington, DC: Mathematical Association of America. Report of an NSF working group convened to support assessment of calculus reform projects by providing a conceptual framework together with extensive examples. Emphasizes the 'fundamental tenet' that, since tests are statements of what is valued, new curricula need new tests.
- Seldin, Peter (1993). The Use and Abuse of Student Ratings of Professors. *Chronicle of Higher Education*, July 21, A40. An op-ed column lamenting the propensity of colleges to misuse student evaluations of faculty. Gives research-based advice for how to use such ratings intelligently and effectively.
- Smith, G., Wood, L., Coupland, M., Stephenson, B., Crawford, K. and Ball, G. (1996). Constructing Mathematical Examinations to Assess a Range of Knowledge and Skills. *Int J Math Edu Sci Technol*, 27, 65-77.
- Steen, Lynn Arthur (1999). Assessing Assessment. Preface to Bonnie Gold, Sanda Z. Keith and William A. Marion (Eds.), *Assessment Practices in College Mathematics*, pp. 1-6. Washington, DC: Mathematical Association of America. An exploration of issues, principles, and options available to address the wide variety of assessment challenges facing college mathematics departments.
- Stevens, Floraline, Lawrenz, Frances and Sharp, Laure (1993). *User-Friendly Handbook for Project Evaluation*. Washington, DC: National Science Foundation. A 'how-to' guide to effective assessment for project directors who have neither experience in nor enthusiasm for evaluation.
- Tucker, Alan C. and Leitzel, James R. C. (1995). *Assessing Calculus Reform Efforts*. Washington DC: Mathematical Association of America. A 'mid-term' review of the NSF-supported calculus reform movement in the United States, providing background on the motivation and goals of the movement, as well as evidence of changes in content, pedagogy, impact on students, faculty, departments, and institutions.
- Wiggins, Grant (1989). A True Test: Toward More Authentic and Equitable Assessment. *Phi Delta Kappa*, May, 703-713. Argues that misunderstanding about the relation of tests to standards impedes progress in educational improvement. Suggests that only tests that require the 'performance of exemplary tasks' can truly monitor students' progress towards educational standards.
- Wiggins, Grant (1990). The Truth May Make You Free, but the Test May Keep You Imprisoned: Toward Assessment Worthy of the Liberal Arts. In *Assessment 1990: Understanding the Implications*, pp.17-

31, Washington, DC: The American Association for Higher Education. (Reprinted in 1992, Lynn A. Steen (Ed.). *Heeding the Call for Change: Suggestions for Curricular Action*, pp. 150-162. Washington, DC: Mathematical Association of America.) Philosophical reflections on the purposes of education in the liberal arts or in basic science or mathematics. Focuses on ten principles of education that testing tends to destroy (e.g., justifying one's opinions; known, clear, public standards and criteria; self-assessment in terms of standards of rational inquiry; challenging authority and asking good questions).

Ken Houston

University of Ulster, Northern Ireland

sk.houston@ulster.ac.uk