# *Assessment of Chemistry Degrees*

PERSPECTIVE

Patrick D Bailey

Department of Chemistry, Heriot-Watt University, Riccarton, Edinburgh, EH14 4AS
p.d.bailey@hw.ac.uk

## The problem

Do we use appropriate methods to award and classify degrees in chemistry in the UK? I fear that, to a large extent, we do not. In part, this is due to us failing to identify appropriate assessment procedures for the full range of skills required by professional chemists. But I believe that external demands for accountability are also undermining the value of standard examinations, by encouraging us to devise highly structured questions for which the marking is transparent, but for which the skills being tested are limited.

I contend that the entirely reasonable demand for accountability has taken us down a track which is to the detriment of university education because it leads to a decrease in real quality of assessment. I agree absolutely with the proposition that Universities should have clearly defined aims and objectives, and that these should be reflected in the procedures used to define the class of degree awarded to students. The problem is that the demand for *transparency* in the process of assessment has been interpreted to mean that the working process must be objective; taken to its limit, this means that the marking process could be carried out by computer. This pressure to achieve objectivity creates three inter-related problems:

- the scope for professional judgement is reduced;
- the range of possible assessment procedures becomes limited;
- opportunity for reflection is destroyed by increasing frequency of examinations.

Assessment should be addressing several issues (e.g. feedback to students, feedback to lecturers, generation of a mark), and the procedures used should be appropriate to the skills assessed. However, most HE teachers risk ignoring the first two and play safe by using traditional unseen test mechanisms which minimise the risk of cheating and which can be marked according to a scheme which involves little or no professional judgement. As a result, most assessment comes close to Ramsden's description of "a parody of bad practice"[1].

## Background

University chemistry educators are committed to running high quality courses; as far as their teaching role is concerned this means quality in provision of a learning environment and quality of the procedure by which student ability is judged. The latter is a crucial factor in determining the future of the individual student. Not surprisingly, departments are under pressure to be able to *demonstrate* the effectiveness of their quality control processes. The pressure comes, for example, from accreditation bodies (the RSC), external assessors (e.g. TQA), influential reports (e.g. Dearing), university administrators (e.g. those responsible for monitoring progression and completion rates), and perhaps also employers. I will argue that the demand that quality should be objectively demonstrated puts an unacceptable limitation on our opportunities to assess quality over the range of abilities which we need to evaluate.

Here are some comments I have collected from colleagues throughout the UK, which support my argument.

- Most examination questions are highly structured questions, so students do not need to make extended arguments, or to write discursive open ended answers.
- Degrees are awarded by accumulating short term, isolated modular credits, rather than general 'professional' skills – and whilst many departments are ensuring that their courses include synoptic components, some universities are now forbidding this!
- The sheer quantity of assessment benefits students who aim to achieve good exam marks by virtue of short term memory, more than those who aim to develop as all round professional chemists.
- University administrators are imposing rigorous classification/mark correlations; thus 59.5% might be a 2:1 (rounded up to 60%), but 59.3% might be a 2:2, even though the error in the marking procedure might be as much as 1 or 2%.

All these comments are consistent with my own observations that most chemistry departments assign around 80% of the total marks available for classification purposes to formal written examinations, and the overwhelming emphasis in these examinations is on structured questions. Even with relatively unstructured questions, mark schemes significantly restrict the opportunity to give credit for anything beyond factual content. Thus, despite pressures on us to create a learning environment in which students develop a wide range of professional skills, we do little to evaluate these skills. Not surprisingly, this affects the way our students develop; students want to score high marks in examinations and they will inevitably focus their attention on those aspects of the course for which we award marks. The structured examination procedure is also convenient for external administrators and assessors; questions can be marked to such a strict protocol that it results in a reliably reproducible mark for each answer. The *reproducibility* of the mark (irrespective of marker) hides the fact that it may be completely unreliable as an indication of the full range of skills we might wish to assess. By limiting our assessment procedures to those aspects

of chemistry that can be marked with high precision (e.g. by an automaton) we assess only a tiny fraction of the attributes we want our students to develop.

I emphasise that I am not arguing that our *marking* is of poor quality; on the contrary, the RSC accreditation process and the quality control exerted by external examiners lead to practice which stands up well to scrutiny. What I do argue is that we are being driven to more rigid *assessment procedures*, and this is to the detriment of our subject.

I identify three factors which should specially influence the assessment process. Firstly, we need to look at what we are expected or required to deliver in any undergraduate degree, and in chemistry degrees in particular. Secondly, we need to identify the opportunities for assessment in typical chemistry degree programmes, and consider how these correlate with course aims. And crucially, we must be aware of how our assessment processes affect the overall development of our students.

## Factors influencing the assessment process

### Chemistry degree programmes

Three recent reports are of special relevance to the design of a chemistry degree programme. These are the Dearing Report[2], the Mason Report[3], and the chemistry Benchmarking Document[4] which is based on the QAA agenda for quality[5].

Recommendation 21 of the Dearing Report sets out in general terms what we should expect to find in any programme on offer at a university.

"We recommend that institutions of higher education begin immediately to develop, for each programme they offer, a 'programme specification' which identifies potential stopping-off points, and gives the intended outcomes of the programme in terms of:

(a) the knowledge and understanding that a student will be expected to have on completion;
(b) key skills: communication, numeracy, the use of information technology, learning how to learn;
(c) cognitive skills, such as an understanding of methodologies or ability in critical analysis;
(d) subject specific skills such as laboratory skills."

The Mason Report, commissioned jointly by the Royal Society of Chemistry and the Council for Industry and Higher Education, examines ways in which the teaching of chemistry in higher education should adapt to employers' needs around the turn of the millennium. Amongst the diverse range of skills that industry would like to see better developed, it is clear that adaptability and communications skills feature highly across all sectors; mainstream chemistry employers are also concerned by limitations in the practical skills of candidates.

The Chemistry Benchmarking Document is a well-balanced document (currently still only consultative), which gives some fairly detailed advice on course content without being too prescriptive; in general, it follows the recommendations made by Dearing, requiring a balance of subject-specific and generic skills, and providing extra detail relating specifically to chemistry. In terms of what we are expected to deliver, it suggests benchmark 'performance criteria', which include aspects of knowledge and understanding, problem-solving, experimental skills, and transferable skills. However, unlike the Law and History panels, who chose to identify only baselines levels for their degrees, the Chemistry benchmarking panel decided to identify a range of standards that undergraduates might achieve in each of four areas identified in Section 6 of the QAA document as "what new graduates should know and be able to do". These four areas, clearly based on recommendation 21 of the Dearing Report, are:

(i) subject knowledge and understanding;
(ii) cognitive skills;
(iii) discipline-related practical and professional skills;
(iv) general transferable skills.

The Benchmarking Document identifies, for each of these four areas, specific attributes corresponding to an A-E classification. Although I applaud the decision, there is a high risk that the A-E classification will be misused. It was never intended to be correlated with a particular class of degree; a student who showed *all* the attributes listed in the A grade would be truly outstanding; for example, I would expect few students achieving (and worthy of) a first class degree to be able to demonstrate knowledge "significantly beyond that covered in the degree programme", although the very best undoubtedly would do so. Unfortunately, I fear that those looking at chemistry degrees from the outside (e.g. for quality assessment or accountability reasons) will use the descriptors for grades A – E precisely as guidelines for the qualities to look for in students in the degree class which apparently corresponds to each category.

In spite of this risk, which we must guard against, the grade descriptors in the benchmarking document give a useful indication of the range of knowledge, understanding and skills which students on chemistry degree courses should have the opportunity to develop. Students with these qualities would meet the criticisms of graduates raised by employers and reported by Mason[3]. However, it is not enough to provide opportunities for students to learn; we must surely also demonstrate our commitment by assessing how well they have learned, and the mark which they obtain must be seen as contributing to their degree class.

### Opportunities for assessment

There are perhaps more opportunities for varied assessment methods in chemistry than in almost any other subject, and the benchmarking panel[3] identified many of them:

"Assessment procedures. It is essential that the procedures used for the assessment of students' achievement in chemistry should correspond to the knowledge, competencies and skills that are to be developed through their degree programme. Evidence on which the assessment of student achievement is based should include:

- Formal examinations, including a significant proportion of unseen examinations.
- Laboratory reports.
- Oral presentations.
- Planning, conduct and reporting of project work.

Additional evidence of use for the assessment of student

achievement may be derived from:

- Essay assignments.
- Problem-solving exercises.
- Portfolios on chemical activities undertaken
- Literature surveys and evaluations.
- Collaborative project work.
- Preparation and displays of 'posters' reporting project work."

The opportunities for variety in assessment does not mean that we actually take advantage of them. I have already argued that there is pressure to move towards formal written examinations with rigid marking schemes, and that this results in most of the alternative procedures playing a rather insignificant part in the process. Key features of learning identified in the Benchmarking document do not map well onto the qualities which are appropriately assessed by formal written examinations. The first two (subject knowledge and understanding, and cognitive skills) can be suitably examined in this way, but it is hard to see how the other two key areas (discipline related practical and professional skills, and general transferable skills) can be evaluated by written tests. In most (probably all) chemistry degree courses, a mark for laboratory work and for project work contributes to the final classification. Both of these provide opportunities to assess various skills, but each typically contributes only about 10% of the total, and the marks are often rather indiscriminatory.

We need to think carefully whether the mix of assessment procedures we actually use to contribute to the classification of degrees reflects both the opportunities created by our teaching, and the skills we wish our students to develop.

**Students**

The assessment procedures we use send important messages to students about what we judge to be the important things for them to learn. Race and Brown[6] suggest ten methods of assessment and, more importantly, identify the advantages and disadvantages associated with each of them, and provide suggestions for making them more effective. They suggest that we should make use of as wide a range of assessment methods as possible, whilst recognising that all have merits and limitations. Unfortunately, they conclude that:

"It can be argued that presently we have far too much assessment, but neither the quality nor the diversity of this assessment is right. Students are highly intelligent people; if we confront them with a game where learning is linked to a rigid monotonous diet of assessment, they will learn according to the rules of that game. To improve their learning, we need to improve our game."

This quotation raises the problem of over-assessment. Students are assessed more extensively than ever before, both to evaluate progress through their chosen course and for the purpose of classifying their final degree. Over-assessment is almost certainly imposed with the best of intentions. A common response to poor examination results is to suggest that *more* tests along the way would help the students by providing early warning and feedback. I suggest that there are three problems with this approach.

First, most of the tests are set in a style which reinforces the student view that learning to be a chemist involves no more than learning the correct answer to a defined range of questions. In this sense, in-course tests do not offer a significant benefit to the learning habit. One of their main advantages should be the provision of detailed feedback. But I know of no evidence to show that the weaker students, who have the greatest need for such feedback, actually benefit from it.

A second problem with continual testing, discussed by Beard and Hartley[7], is that it takes little account of the different learning characteristics of students, some of whom appreciate "working gently through the year", while others "think continuous assessment is more strain". In the same survey, another student commented that the regular revising for tests prevented extra work like "background reading".

This quotation raises the third problem of over-assessment: the impact on the time available for reflection. This is an essential feature of in-depth learning. For example, Johnstone[8] has developed some highly successful new teaching material, the design of which was guided by the following model:

(a) the learning process uses working space, which is fed by external events/observations/instructions;
(b) the information that we select from the external input is controlled by a filter mechanism, which uses knowledge that is stored in our long term memory;
(c) working memory is only really effective when we have the chance to order our thoughts and seek out the inter-relation between various pieces of information;
(d) steps a – c are *all required* if we are to transfer information usefully from our working memory to our long term memory.

Without the opportunity for reflection, the long-term memory can simply become a jumble of unconnected facts, or information is largely retained in the short-term memory. This situation corresponds alarmingly well with the criticisms often levelled against many students who appear to forget topics once they have been assessed.

These three disadvantages of over-examining do not mean that there is no room for assessment at appropriate stages through the course. However, the prime purpose of this course-assessment (or continuous assessment) needs to be to encourage learning and provide feedback for both tutors and students. Learning is promoted by encouraging students to reflect on new ideas and incorporate them usefully into their long term memory. Feedback occurs when both tutors and students are able to recognise and rectify their own shortcomings (in delivery or in learning). For practical purposes, it may also be important to assign a numerical mark to these assessment procedures because it is this which convinces students that we treat them seriously.

The problem is that assessment methods which encourage learning and provide useful feedback are not usually totally objective but involve a degree of professional and subjective judgement. Students are inclined to regard this as 'unfair' because it is not immediately obvious how the mark was arrived at. Many students find it difficult to understand why good work (which is perhaps 80% correct in terms of factual content) may be worth a mark of only 60%. They are apparently unaware that it is the quality of presentation,

additional knowledge, or other subtle observations, that make all the difference between good work and excellent work.This is true in the real world. It is perfectly reasonable that the same criteria should be applied to our assessments at university. Persig's observations of students' ability to identify and appreciate subjective quality[9] should encourage us to have the courage to make and defend professional judgements.

One possible way of introducing variety into assessment methods in a way which enhances learning is to make use of peer marking. Some success has been reported with various approaches to this[10,11], though other reported problems show how important it is for the students to believe that the process is well thought out and fair[12]. I suggest that, as an extension to the principle of individual students marking an individual piece of work, there is room for students to work (occasionally) in groups to mark and rank five or six pieces of work from colleagues. This might help them to appreciate the criteria we use to judge a piece of work and to recognise the difference between good work (e.g. 80% correct but worth only a mark of 60%) and excellent work worth a first class mark. This would help them to understand that there is a difference between *transparency of process* (the process by which we assign a mark) and strict allocation of each available mark to a particular piece of information. Even in a formal written examination it is possible to use a marking scheme which includes some marks for the skills demonstrated in applying and presenting knowledge. This is bound to be at least to some extent subjective and it would benefit both our students and our profession.

## Conclusion

With the new QAA accreditation process now being trialled, and the provisional chemistry bench-marking document in place, we have the opportunity to address these assessment issues. We can allow the problems to get worse or make determined efforts to improve our assessment. Here are my suggestions for improvement.

- We should aim to integrate our assessment procedures more fully into the learning process, and thus emphasise the importance of feedback and self-assessment. Ramsden[1] suggests ways in which this might be done, and Rowntree[13] has written an excellent book on this.
- The assessment procedures we use need to match the knowledge and skills we wish to assess, and we simply have to allocate marks explicitly for skills if we want students to take them seriously. As Hartley and Braithwaite point out[14], is it any wonder that students gear their work specifically towards tests and exams, when this is precisely where we allocate their marks?
- We should not be afraid to use our professional judgement in assessing skills which do not lend themselves to objective measurement.We do this in other aspects of our work, and students need to appreciate that it is a feature of society.In doing this we have to recognise that *transparency of process* does not always imply *objectivity of marking*.
- We should aim to decrease the amount of assessment, whilst increasing the variety of methods used. In particular we should look as carefully at the cumulative assessment process over the whole course as we look at the assessment of each unit or module. For example, it may be that one can justify the assessment methods for each module, whereas there may be a clear over-examining of students when the course is viewed as a whole. A reduction in the amount of assessment might help students reflect on their work and gain a better understanding of it; and focussing on assessment methods that are appropriate for different skills would surely help our graduates to become more rounded professional chemists.

In summary, most chemistry courses do seem to address the 'programme specification' outlined by Dearing, and detailed in the QAA template. Most chemistry courses are RSC accredited, meet the guidelines of the benchmarking panel, and are robustly assessed. But are our graduates achieving the full range of professional skills that we, and employers, would like? And are our assessment procedures really encouraging our students to develop these skills?

## *References*


1. Ramsden P,*Learning to Teach in Higher Education*, Routledge, London, 1992, 181-213.
2. Dearing R, 1997,*Higher education in the learning society*, Report of The National Committee of Enquiry into Higher Education (Chairman), HMSO.
3. Mason, G, 1998, Report on *Change and Diversity: The challenges facing chemistry higher education*,RSC.
4. *The Bulletin of the Quality Assurance Agency for Higher Education*, **1(3)**, March 1998.
5. *General Guidelines for the Academic Review of Bachelors Honours Degree Courses in Chemistry*, Quality Assurance Agency for Higher Education, 1998.
6. Race P and Brown S, 1998, *The Lecturer's Toolkit*, Kogan Page, London, 51-78.
7. Beard R and Hartley J, 1984, *Teaching and Learning in Higher Education* (4th edition), (Part 5), Harper & Row, London, 273-296.
8. Johnstone A H, 1997, *UChemEd*, **1**, 8-13;p Johnstone A H, 1997, *JChemEd*, **74**, 262-268.
9. Persig R M, 1974, *Zen and the Art of Motorcycle Maintenance*, 208-213, Vintage.
10. Murray R, 1998, How to mark 100 essays in 2 hours, *Proceedings of Variety in Chemistry Teaching* (Garratt J and Overton T eds) RSC.
11. Duckett S B, Lowe N D, Taylor PC, 1998, Addressing Key Skills in the Chemistry Curriculum, *UChemEd*, **2**, 45-50.
12. Clow D J M, 1996, Computer Simulations of laboratory experiments, D.Phil Thesis, University of York.
13. Rowntree D, *Assessing Students*, Harper & Row, London, 1977.
14. Hartley J and Braithwaite J A, *Durham Research Review*, **8**, 14-20 (see also ref. 7, 275).