Research article

The Findings of an Assessment Audit: an NTFS Project Report

Ian Hughes

Professor of Pharmacology Education, Bioscience Centre, Higher Education Academy, School of Biomedical Sciences, University of Leeds, Leeds, LS2 9JT, UK

Date received: 23/08/05 Date accepted: 21/12/05

Abstract

An Assessment Audit is described consisting of 47 questions, each being scored 0 to 4, by the module team depending on the extent to which the audit point was satisfied. Scores of 2 or less indicated unsatisfactory provision. Audits were carried out on 14 bioscience- or medicine- based modules in 13 universities.

There was great variability between modules in the hours spent by:

- *teachers in direct contact with students (12 to 914);*
- *teachers involved in the process of assessment (2 to 372);*
- *students actually being assessed (2 to 60 hours per student);*
- *students in the teaching and learning process (35 to 300).*

The highest scoring module obtained 133 out of 188 (71%) with 11 out of 47 items scored at 2 or less while the lowest scoring module obtained 47% with 27 items scoring 2 or less. Features consistently poorly addressed were:

- *consideration of learning objectives/assessment in other modules taken by the student;*
- *consideration of consistency among multiple markers;*
- *use of known mark sets to validate data processing;*
- *availability of exemplar answers;*
- *feedback on end-of-module assessments.*

A common issue concerns the isolation of modules and module teams. This suggests the need for a strengthening of the course thread and emphasis on the totality of the student learning experience rather than the individual module. The audit provides a framework within which course teams can reflect on and improve the quality of the assessment in their module.

Keywords: Assessment, assessment audit, quality of assessment, improving assessment, improving assessment through audit.

Introduction and aims

Assessment of students occupies a great deal of staff and student time and is often a contentious area, frequently giving rise to complaints from students. Two areas in particular give rise to complaints, these being the arrangement

of assessment in the course and the timing, amount or quality of the feedback available after the assessment has been returned. Because of the importance of assessment in courses it seemed appropriate to develop an Assessment Audit to encourage teachers to re-appraise the quality of the assessment procedures on their modules and to think, reflectively, about how they could be improved. There have been a variety of other initiative to try to improve assessment practices both by individuals (Palomba & Banta, 1999; Wakeford, 1999) and organisations such as the (then) LTSN Generic Centre (Smith, 2004) and the Quality Assurance Agency (2000).

Design criteria

I devised the Assessment Audit over a period of about 6 months and then trialled it with several colleagues to whom I am very grateful for comments. There was a need to keep it relatively short (so busy teachers would actually use it) and located at the coal-face (module teams) so its results would be accepted and that local ownership would help changes to be implemented (Jones 1997). It also needed to be comprehensive and to include the important peripheral issues often raised by students and staff and comprehensively reviewed elsewhere (Heywood, 2000). I decided initially to develop a paper-based audit as I would normally expect it to be completed by the module team sitting round a table. In addition, this gave me more control over getting responses from the trailers. The version of the audit used in 2001 is included here as appendix A (note: this appendix also includes, for each audit item, the percentage of the maximal possible score achieved; see *General Results* below). The current version, which has evolved somewhat, can be downloaded from http://www.bioscience.heacademy.ac.uk

Although the audit was intended for use formatively rather than summatively each audit question was required to be scored as it was thought that this would concentrate minds on whether the current provision of assessment within a module was ideal or could be improved. It also allowed for the identification and prioritisation of particular aspects that were poorly addressed and allowed for aggregation of results from several different modules or universities. The aggregated results from 14 audits (from 13 separate universities) are reported here.

Methods

'Volunteer' academics involved in the running of at least one module were recruited into the trial and ran the Assessment Audit (with their module management teams) for the module under consideration. It was agreed that on completion of all aspects of the audit their department or unit would be paid a small sum to acknowledge the time involved. Time frames for delivery of the audit were agreed and set out in a letter agreed by each participant. One participant withdrew early because of pressure of other work, but all others completed the required process. Participants were asked to select a module in which the assessment processes were less than ideal to be subjected to audit. The universities of Glasgow, Manchester Metropolitan, Wolverhampton, Leeds, Birmingham, Keele, Monash (Australia), Birkbeck London, Westminster, Kent, Dundee, Cardiff and Bath Spa University College took part. All modules were part of preclinical medical courses or biological or

biomedical science degrees. In the 8 aspects of assessment which comprised the Assessment Audit participants were asked to score 0 to 4 for **EACH** of the 47 INDIVIDUAL audit points depending on how closely these audit points are **optimally** achieved. The instructions were to score 0 if the audit point has not been serious considered at all and to score 1-4 if the audit point has been considered but reflection indicates that it is poorly (1), partially (2), adequately (3) or optimally (4) satisfied. Not every audit point was applicable to every module and some data therefore derive from less than the complete sample of 14 audits.

Some numeric data were requested at the end of the audit particularly on the hours involved in teaching and assessment by staff and by students. These data were not fully complete, showed wide variability and were not normally distributed. They were therefore not readily or meaningfully amenable to statistical treatment. Where appropriate, the crude data sets have been presented.

The audit also asked for a commentary on how audit items scoring 2 or less could be addressed. Audits were completed between March 2002 and June 2003 and returned to Leeds where the data from the 14 audits was processed in a tabular database.

Results and findings of audits

A. General results

First, second and third year undergraduate modules were included. The highest scoring module obtained 133 out of a possible 188 (71%) with 11 out of 47 items being scored at 2 or less while the lowest scoring module obtained 47% with 27 of the 47 items scoring 2 or less. The actual totals were: 113, 90, 114, 89, 132, 104, 112, 128, 133, 91, 111, 99, 104, 121 and it should be noted that the values in the data sets throughout this paper are always presented in the same order with respect to the modules to which they pertain. There was therefore considerable variation in the quality of the assessment in the audited modules as measured by the Assessment Audit.

The scores returned for the individual audit points have been used to calculate a mean score for each item and these have been expressed as a percentage of the maximum possible total score for the item and are shown in Appendix A.

B. Items generally well fulfilled.

Some mean scores showed there had been a generally good performance with respect to particular items in the audit. Items scoring more than 70% and therefore satisfactorily fulfilled by most respondents are shown in Table 1. Although these represented only 12 out of the 47 audit items it must be remembered these modules were selected because it was known their associated assessment was not optimal.

Table 1 *Showing those audit points generally well satisfied (i.e. scoring, on average, more than 70% of the maximum score).*

Assessment Audit Item	**Score (%)**
1.1 Are the learning objectives (i.e. the changes in the student's knowledge, skills and attitudes) explicit for the module and for each constituent element piece of work where appropriate?	73
1.7 A single type of assessment (e.g. all MCQs) may disadvantage some students. Is there a variety of assessment methods used in different circumstances?	73
2.6 Are the assessment methods appropriate to the teaching style used?	77
3.1 Are there grade descriptors available to the students?	76
3.2 Are these (grade descriptors) known to and followed by the staff doing the marking?	83
3.5 Are the grade descriptors congruent with those on other modules taken by the students?	80
4.1 Is the mark distribution for each piece of work known and considered?	71
4.4 Is there external moderation of the marks	75
6.1 Is feedback provided on all in course assessments?	80
6.3 Is feedback provided to all students?	86
7.1 Are resit arrangements written, available to the students and explicit with regard to format and material covered?	80
7.3 Are the Learning Objectives the same in resit examinations?	80

C. Items generally poorly fulfilled.

The more interesting data are related to items in the 14 audits that were consistently given scores of 2 or less. Table 2 presents the items that were scored 2 or less in 7 or more (50% or more) of the 14 modules audited. Comments on each item are included below.

Table 2 *Showing the Assessment Audit items for which the majority of modules were found to be poorly satisfying the requirements.*

Assessment Audit Item	**Number of unsatisfactory scores (2 or less) / total number responding**
1.2 Are the different types of element in the Learning Objectives reflected in the assessment*? (e.g. knowledge, understanding, skills, attitudes etc).*	7/14
1.3 In setting the Learning Objectives is consideration given to the learning objectives in other concurrent or previous modules?	11/13*
1.4 Is the different achievement in each Learning Objectives *separately* identifiable by the student in the overall assessment?	9/14
2.3 If multiple markers are used is uniformity of marking tested and, if necessary, compensated for?	6/9$
2.4 If double marked is there a mechanism other than taking the average to resolve significant differences?	7/14
2.8 Are known mark sets included in the mark spread sheets to demonstrate accuracy of mathematical processing/combining of marks?	11/14
2.9 Is there external input to the assessment process?	7/14
3.3 Are there exemplar answers?	10/14
3.4 Are exemplar answers available at different grades?	14/14
5.1 Does assessment provide a monitor of student performance throughout the module?	9/14
5.2 Is there time to allow students to respond to a poor assessment before the end of the module?	10/14
5.5 Is assessment timely with regard to: other assessment on other modules?	11/14
6.2 Is feedback provided on the end of module assessment?	13/14
6.6 Do you know that all students access the feedback provided?	9/14
6.7 Are students performing poorly counseled (on a one-to-one basis)?	9/14
6.8 Does counseling take into account performance on other modules?	9/14
7.2 Is the date/time of any resit exams known to the students at least 3 months before it takes place?	10/14
7.4 Are resit candidates given effective feedback on their performance in the first sit?	9/14

** - not applicable to one module; $ - only applicable to some modules*

1.2 Are the different types of element in the Learning Objectives reflected in the assessment? (e.g. knowledge, understanding, skills, attitudes etc).

Clearly, assessment is not well matched to the learning objectives. As a good example, assessed practical write-ups are hardly a reasonable measure of skills in the laboratory, hence the development of lab skills, often cited as a learning objective, is rarely assessed. It is very difficult to assess an individual's laboratory skills (outside the final year practical project) as many laboratory classes are performed in groups and/or with large numbers. Peer assessment is one possible solution. Another is requiring an exercise with a known quantitative answer or designing the class so half the student create a product and the other half analyse it.

1.3 In setting the Learning Objectives is consideration given to the learning objectives in other concurrent or previous modules?

The diversity of module choices available to students makes this difficult to achieve and the provision of an alternative for students who have already achieved a required learning objective can be resource intensive. In terms of skills, the use of a skills matrix, completed by students, can help identify areas of duplication but repetition may be useful especially for weaker students. Clear definition of course learning objectives and consideration of where these are achieved may be useful in identifying duplication or gaps resulting from particular module choices made by students. The small number of modules (2) on which this audit point was well satisfied emphasises the isolation of each module and the lack of a strong course/programme strand.

1.4 Is the different achievement in each Learning Objectives separately identifiable by the student in the overall assessment?

To learn that you have achieved overall 50% does little to help a student identify which areas need improvement. If feedback is to lead to remedial action by students then the nature of the weakness must be identified. This emphasises the need for quality, specific feedback related to each of the learning objectives. The development of personal portfolios may go some way to address this issue. However, having separate scoring of knowledge, analysis and critical approach for example within the assessment of a piece of work, e.g. an essay, is a new concept for some staff members and there is some staff development work to be done in this area.

2.3 If multiple markers are used is uniformity of marking tested and, if necessary, compensated for?

Differences might arise not just because students are marked differently but also because students are taught differently (e.g. on split sites). Both are important issues. This has also come to the fore with the increasing use of postgraduate students both as teachers and as assessors. With small groups of say 10 students it is not easy to decide what variability in mean mark between groups might be expected if teaching and assessment were uniform. Appropriate training for postgraduates participating in teaching and assessment is essential. The provision of detailed marking proformas is very helpful. Rotating tutors through tutorial groups has the disadvantage that development of relationships is more difficult but it does spread differences in

standards and abilities more evenly and also allows statistical analysis of tutor marking to identify more easily differences in standards.

2.4 If double marked is there a mechanism other than taking the average to resolve significant differences?

The easy route is to average the two marks and this is often done if the difference of 5% or less. Where the difference is greater than this averaging the two marks does not make optimal use of double marking since one of the two marks may simply be an error by the marker. Some universities utilise a third marker while others get the two markers to produce an agreed mark. Some universities adjust the marks of both markers to the same mean value before resolving differences. Some get the external examiner to resolve the difference.

2.8 Are known mark sets included in the mark spread sheets to demonstrate accuracy of mathematical processing/combining of marks?

The problem is compounded by the increasing use of spreadsheets containing multiple entries and complex weighting formulae. The difficulty is illustrated by two situations. First, where the program worked out all the marks correctly and then, craftily, assigned the overall mark to the student NEXT in the list to the one who should have received it. Second, where a spreadsheet correctly worked out the marks except on the line following a student who had withdrawn, and therefore had an incomplete record, where an incorrect overall mark resulted. Inclusion of 4 fictitious students with marks of 25, 50, 75 and 100% in each component gives confidence but does not guarantee there is no problem. Increasing workloads and time pressures have increased the likelihood that errors in processing marks go undetected.

2.9 Is there external input to the assessment process?

All UK universities use external examiners for final year assessment and it is becoming increasingly common to have an external examiner for earlier years as well. External verification of the assessment process for all years provides an independent and robust defense against increasingly common appeals against assessment outcomes. It is usual for different external examiners to have responsibility for final year and for earlier years.

3.3 Are there exemplar answers?

3.4 Are exemplar answers available at different grades?

It is possible to address this issue without large resource implications by providing suitably anonymised student exam answers in each grade. The students' permission must of course be obtained but this is rarely refused. Secretarial staff (providing such resources are available) can transcribe the answer in order that they can be made available on an intranet. This not only provides grade exemplars as feedback for the students who have just sat the exam but also provides model answers for the students who will take the course next year. It is probably best to make available only the last set of examination question answers and not to use an identical question for at least three years. It should be noted that this is not a complete solution since some of my data (unpublished) suggest that only the better students make use of this facility.

5.1 Does assessment provide a monitor of student performance throughout the module? and
5.2 Is there time to allow students to respond to a poor assessment before the end of the module? and
5.5 Is assessment timely with regard to: other assessment on other modules?
[So the totality of the assessment as experienced by the students is reasonably distributed and does not all take place in an unreasonably short period.]
This is a crucial issue and high on the list of student concerns. With increasingly large classes and use of postgraduate student demonstrators/tutors, assessed performance may be the ONLY way academic staff can become alerted to a student with problems. Attendance is a poor proxy for performance but there are systems being developed which allow monitoring of student attendance at lectures using a bar-code system (Gent, 2003). Widening participation and retention issues will make this an even more important area. The very poor performance scores on 5.5 where only 3 of 14 modules were satisfactory again emphasise the isolation of each module and the lack of a strong course/programme theme.

6.2 Is feedback provided on the end of module assessment? and
6.6 Do you know that all students access the feedback provided? and
6.7 Are students performing poorly counseled (on a one-to-one basis)?
This is a crucial issue, high on the list of student concerns, and will become still more important as widening participation and retention issues come to the fore. The poor performance in 6.2 may reflect the difficult of contacting students once a module is complete particularly if it is a second semester module. The use of exemplar answers can be helpful (see 3.3/3.4 above). It can be very beneficial to discuss in detail strengths and weaknesses of student final year projects but this is very resource intensive. The use of a marking sheet on which a marker can write comments is useful but staff need to be educated in what are acceptable and unacceptable comments. For in-course work, systems are available which allow staff to construct an individual feedback report (which is then emailed automatically to the student) by assembling selected pre-typed comments and corrections. This greatly reduces the staff effort required to provide individual feedback (Pitts & Bolton, 2004).

6.8 Does counseling take into account performance on other modules?
The small number compliant again emphasises the isolation of each module and the lack of a strong course/programme strand. While the module manager and their team may be involved only with a particular module the student's experience is a sum of the modules taken. Suitable administrative arrangements allow this to be accomplished through the personal tutor system.

7.2 Is the date/time of any resit exams known to the students at least 3 months before it takes place?
There are increasing problems with students having arranged holidays or jobs or (if overseas) booked return flights up to 9 months in advance. There is a concern as to the extent to which the University might be liable for part of the

cost of any cancellation necessary. Even if the exact dates are not known it is advisable to publish at the start of each year the dates of the 7-10 day resit period and make it clear students may be required to be present in the University during this period.

7.4 Are resit candidates given effective feedback on their performance in the first sit?

See sections 3 and 6 above.

Overview of above results

It is clear from the foregoing that a variety of items in the audit are consistently poorly addressed but there are possible solutions to these problems which are realistic and usable in various contexts. One general theme which appears strongly is that problems arise from the independence of modules and lack of communication between module leaders, each of whom tends to treat their module as an isolated entity. This is perhaps understandable in the environment in which module leaders operate but does point to the need for:

- a **strengthening of the course thread** and
- an emphasis on **the totality of the student learning experience considered in a longitudinal context** rather than at the individual separate module level.

D. Analysis of numerical data elicited in audit.

Numeric data were requested concerning 4 items:

a) Hours spent by teachers (including demonstrators, PG tutors etc) on assessment (…hrs)

b) Hours of direct teacher contact with students (…hrs)

c) Hours spent by students being assessed (…hrs)

d) Total hours that each student was involved in all forms of teaching/learning, directed and self-directed (…hrs)

While these questions appeared to be reasonably clear and did not give problems in the initial informal trials, the variation in the numerical returns gave cause for concern that there might be differences in interpretation by the academic staff running the audits. The staff members responsible for some of the audits were therefore telephoned to ensure that equivalent interpretations of the questions were being made in every case. Some returns were revised in the light of these discussions and the data below reflect these updated results. Questions (c.) and (d.) do contain a subjective element based on staff and student estimates of hours involved (for example, in preparing an assessed practical report or the time involved in self-directed elements) and are therefore the least certain of the data.

As shown in Table 3 there were large variations in the hours spent on the modules:

- by teachers in direct contact with students (12 to 914 hours);
- by teachers involved in the process of assessment (2 to 372 hours);
- by students actually being assessed (2 to 60 hours per student);

- and by students in the teaching and learning process (35 to 300 hours).

These data will be influenced by the numbers of students on the course, by the teaching and assessment styles adopted and by the credit rating of the module.

Hours of direct teacher contact with students varied from 12 to 914 hours (76-fold range) but these were extreme values and the other 12 modules fell in the range 22 to 149 hours (6.8-fold range).

The **hours spent by teachers on assessing** a module varied from 2 to 372 hours (186-fold range) These were extreme values however and the other 12 modules fell within the range 30 to 120 hours (4-fold range).

The high values are NOT necessarily associated with modules with large numbers of students. When converted to minutes of staff time spent on assessment PER STUDENT (Table 3), values ranged from 0.54 minutes to 754 minutes (1396-fold range). Eliminating the two extremes, the remaining values were between 16.1 and 192 minutes/student (12.3-fold range). Calculating assessment time expenditure on a per student basis has therefore not reduced the variability between modules.

Taking a ratio between hours spent in assessment / hours spent on teaching gave values of 2.1, 0.4, 0.03, 1.2, 0.5, 1.1, 7.3, 2.7, 0.9, 0.5, 0.9, 0.4, 1.4, 3.9. It is interesting that on 7 modules the ratio exceeded unity indicating teachers spent more time assessing the students than they did in direct contact providing teaching. In four cases, more than twice as much time was spent in assessment than in direct teaching.

The hours spent by each student in being assessed (Table 3) varied from 2 to 60 hours (range 20-fold). The higher figures were associated with modules that involved practical work and required assessed write-ups, but the previous comments about the reliability of these data must be borne in mind.

The total hours spent by a student on a module in all forms of teaching or learning activities varied from 35 to 300 hours (range 8.6-fold). As noted above these values are associated with some uncertainty as they involved student or staff estimates of hours involved in self-directed study.

Table 3 *Showing the student numbers and the time spent by staff and students in teaching, leaning and assessment for each of the 14 modules audited. The range of reported values is also shown.*

Audit measure	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**Range of values**
Student numbers	100	340	220	80	65	130	7	54	20	120	50	10	#	50	
Staff hrs assessment	46	372	2	88	54	35	88	108	43	78	30	32	60	120	186
Staff hrs teaching	22	914	60	75	117	30	12	40	48	149	33	75	42	31	76
Student hrs assessment	3	‡	2	60	7	5	3	11	5.5	8	44	60	6	‡	30
Student hours of learning	200	180	300	38	35	36	146	200	120	55	300	‡	150	‡	9
Staff MINUTES assessment per student	28	66	0.54	66	50	16	754	120	129	39	36	192	#	144	1396
Staff MINUTES teaching per student	13	162	16	56	108	14	102	44	144	74	40	450	#	37	12

 # – variable in different parts of module; ‡ – depends on options chosen

Effect of credit rating of module

Not all modules subjected to assessment audit carried the same credit rating and some were not credit rated at all. To examine the extent to which different credit rating is responsible for the variation between the modules, the values have been expressed where possible on a ‘per credit’ basis (Table 4).

Even when the credit rating of the module is allowed for the variation between modules is still considerable as it is when both credit rating and student numbers are allowed for (Table 4).

Table 4 *Showing the effect of allowing for the credit rating of the module on the variability of time spent by staff and students in teaching, learning and assessment for each of the 14 modules audited. The range of reported values is also shown.*

Audit measure	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**Range of values**
Credit rating	10	20	20	30	15	10	*	15	20	6#	*	30	*	15	
Staff hrs assessment per unit credit	4.6	18.6	0.1	2.9	3.6	3.5	*	7.2	2.1	#	*	1.1	*	8.0	186
Staff hrs teaching per unit credit	2.2	45.7	3.0	2.5	7.8	3.0	*	2.6	2.4	#	*	2.5	*	2.1	22
Student hrs assessment per unit credit	0.3	†	0.1	2.0	0.5	0.5	*	0.7	0.3	#	*	2.0	*	‡	20
Student hours of learning per unit credit	20.0	9.0	15.0	1.2	2.3	3.6	*	13.3	6.0	#	*	‡	*	‡	16
Staff MINUTES assessment per unit credit per student	3	3	0.03	2	3	2	*	8	6	#	*	11	*‡	16	355
Staff MINUTES teaching per unit credit per student	1	8	0.8	2	7	1	*	3	7	#	*	15	*‡	3	18

** - not credit rated; # - Australian credit rating system not necessarily comparable to UK system; † - depends on options chosen; ‡ - different student numbers in different parts of module*

Overview of above results

Even given the uncertainty of some of the data some conclusions are clear.

- Modules involve very different time commitments from staff with regard to assessment of the students.
- In some cases more staff time was spent on assessing the students than on teaching them.
- The hours of student work involved in a module was very variable.
- The time spend by a student being assessed was very variable between modules.

Given the differences in content, level and numbers of students on the modules, it is unlikely that there are any optimal answers or targets that are appropriate. Teaching styles will affect the data as, for example, it may take a great deal more staff time to teach 360 students in groups of 5 than it does to teach them as a single group. The type of assessment used will also affect the data since it takes much more staff time to mark essay-type answers than to process machine-readable MCQ assessments. This also of course raises a quality issue and an issue of appropriateness of assessment regarding the learning objectives of the module. In auditing the assessment of a module, it is nevertheless instructive to calculate the above data and, in line with the

formative purpose of the assessment audit, to think if the staff and student time occupied is being used optimally. Clearly, some modules are run with remarkably little staff time being spent on assessment.

Discussion

University staff interested in participating in the project were asked to audit a module for which the assessment was already known to be in need of improvement. It cannot be assumed therefore that the results presented below are, in general, representative of the assessment on taught modules in universities generally.

Nevertheless, the findings show that:

1. Several features are consistently poorly addressed with regard to assessment on modules. For example:

- consideration of work/assessment in other modules taken by the student;
- consideration of consistency among multiple markers;
- use of known mark sets to validate data processing;
- availability of exemplar answers; and
- feedback on end-of-module assessment.

2. Despite inevitable subjectivity by both staff and students in their estimations, there is nonetheless great variability in the time spent in both teaching and assessing individual modules

3. There is a need for consideration of the longitudinal, holistic view of the student learning experience rather than the current focus on the individual module.

From my own perspective, I have enjoyed carrying out this project; it has been fun, stimulating and I have been involved with some interesting people. It has broadened my knowledge of assessment and made me think about assessment generally and about my own practice in particular. Many of those who used the audit commented that it had been an extremely useful exercise and that it had caused them to change their practice significantly. Several went on to apply the assessment audit to other modules in their institution. Disseminating the audit through the Bioscience Centre, Higher Education Academy has enabled the assessment audit to be used by a much wider range of teachers than would otherwise have been possible both within the area of Bioscience and outside it.

The success of the assessment audit and its applicability for use in a workshop context has generated a realisation that the same methodology can be applied to other areas of teaching practice. Hence other audits have now been developed, including:

- an Employability Audit for use at a course level (Macfarlane-Dick, 2005);
- a Placements Audit;
- a Skills Audit;

- a Work-related Learning Audit;
- for undergraduate courses, an Examiners External Examiners Audit and its counterpart for institutions (Institutional External Examiners Audit); and
- an Ethics Audit.

All these are available through the Bioscience Centre, Higher Education Academy.

Acknowledgements

Financial support from the National Teaching Fellowship scheme is gratefully acknowledged as is the help of those who contributed to the audits in their institutions.

Communicating author Ian Hughes, Professor of Pharmacology Education, Learning and Teaching Support Network Centre for Bioscience, University of Leeds, Leeds, LS2 9JT, UK: Tel 0113-343-4313; Fax 0113-343-4228; i.e.hughes@leeds.ac.uk

References

Gent, J.P. (2003) Using student barcode information for monitoring student attendance in large groups. http://bio.ltsn.ac.uk/projects/tdf/gent.htm (accessed 28 June 2004)

Jones, S. (1997) *Developing a Learning Culture.* Maidenhead, UK: McGraw Hill

Heywood, J. (2000) Assessment in Higher Education; Student learning, Teaching, Programmes and Institutions. London, UK: Jessica Kingsley

Macfarlane-Dick, D. (2005) Employability Audit. http://www.ltsn.ac.uk/genericcentre/index.asp?id=20473) (accessed 28 June 2004)

Palomba, C.A. and Banta, T.W. (1999) Assessment essentials: Planning, implementing and improving assessment in higher education. San Francisco, USA: Jossey-Bass

Pitts, C. and Bolton, M. (2004) *Automating feedback to students.* http://www.ldu.leeds.ac.uk/ldu/projects/cit/cit7.htm (accessed 1 July 2004)

Quality Assurance Agency (2000) Code *of practice: Section 6: Assessment of students.* http://www.qaa.ac.uk/academicinfrastructure/codeOfPractice/section6/default.asp (accessed 22 December 2005)

Smith, B. (2004) Principles of Assessment*. LTSN Bioscience Bulletin*, **12**, 10

Wakeford, R. (1999) Principles of Assessment*. In: A handbook for teaching and learning in higher education*. Ed. Fry, H., Ketteridge, S., & Marshall, S. London, UK: Kogan Page pp 58-69

Appendix A

Assessment audit tool (working version Oct 2001)

(Note: this appendix also includes, for each audit item, the percentage of the maximal possible score achieved; see *General Results* above)

The purpose of this audit is developmental, not simply to come up with an overall score for the module. It is designed to help teachers **consider** the assessment processes in their module in order to **identify areas** where assessment could be **improved**.

To prepare for the audit it is best to identify each instance when there is assessment in the module and to list for each instance the method(s) used, the students time involved and the staff time involved in marking.

Having done this carry out the audit below with respect to the totality of assessment. The numeric scoring system allows for the situation where, say, against a criterion of 'is marking anonymous?' some (e.g. essays written under exam conditions) meet the criterion and some (e.g. assessed practical write-ups) do not since a score of 1-4 could be recorded depending on the extent of compliance.

In the 8 aspects of assessment score (0-4) for **EACH** of the characteristics of good assessment depending on how closely you believe these characteristics are achieved in the module being audited.
Score 0 if the characteristic has not been serious considered at all.
Score 1-4 if the characteristic has been considered but reflection indicates that it is poorly (1), marginally (2), adequately (3) or completely (4) satisfied. Make a subjective judgement and score accordingly. Note that students can be over as well as under assessed and that assessment can occupy excessive amounts of staff time.

Are the assessment methods appropriate to the learning objectives?

1.1 Are the learning objectives (i.e. the changes in the student's knowledge, skills and attitudes) explicit for the module and for each constituent element piece of work where appropriate? **73%**
1.2 Are the different types of element in the Learning Objectives reflected in the assessment*? (e.g. knowledge, understanding, skills, attitudes etc)* **68%**
1.3 In setting the Learning Objectives is consideration given to the learning objectives in other concurrent or previous modules? **34%**
1.4 Is the different achievement in each Learning Objectives **separately** identifiable by the student in the overall assessment? **43%**
1.5 Are assessment methods/conditions adjusted appropriately for disabled students? **54%**
1.6 Do students experience the method of assessment before it is used summatively*? Either in this module or in a previous module?* **62%**
1.7 A single type of assessment (e.g. all MCQs) may disadvantage some students. Is there a variety of assessment methods used in different circumstances? **73%**

[For example knowledge can be assesses using MCQ, EMQ, SAQ, essays marked for factual content etc. To what extent are different assessment techniques use to give the student a variety of ways in which to demonstrate their abilities?]

The assessment methods used - are they known to provide a secure assessment appropriate to the teaching style?

2.1 To what extent are the methods subjective? **64%**
2.2 Are assessments made from written and agreed marking schemes? **67%**
2.3 If multiple markers are used is uniformity of marking tested and, if necessary, compensated for? **42%**
2.4 If double marked is there a mechanism OTHER than taking the average to resolve significant differences? **58%**
2.5 Is marking done anonymously? *[If machine marked score 4]* **66%**
2.6 Are the assessment methods appropriate to the teaching style used? **77%**
[e.g. if the course is primarily taught using problem-based-learning it would be INappropriate for the assessment to be wholly based on MCQ designed to test factual knowledge].
2.7 Are the students clear as to what would be deemed to constitute plagiarism and has assessment been designed to discourage/prevent plagiarism? **66%**
2.8 Are known mark sets included in the mark spread sheets to demonstrate accuracy of mathematical processing/combining of marks? **36%**
2.9 Is there **external** input into the assessment process? **57%**

Are there published marking criteria and grade descriptors available to the student?

3.1 Are there grade descriptors available to the students? **76%**
3.2 Are these known to and followed by the staff doing the marking? **83%**
3.3 Are there exemplar answers? **41%**
3.4 Are exemplar answers available at different grades? **16%**
3.5 Are the grade descriptors congruent with those on other modules taken by the students? **80%**

How is pass mark decided?

Peer or criterion referenced? [Peer referenced is here defined as the pass mark/grade boundaries being defined in the light of the actual achievement of the student body as a whole. Criterion referenced is where these are decided independently of the actual achievement of the student body.] If peer referenced divide score for this complete aspect by 2.
4.1 Is the mark distribution for each piece of work known and considered? **71%**
4.2 Is the distribution of marks in the module compared with that of previous year's cohorts? **59%**
4.3 Is data available and used to compare the distribution of marks of a student cohort in this module with that in other concurrent modules? **59%**
4.4 Is there external moderation of the marks **75%**
4.5 IF PEER referenced: are the grade boundaries set by a standard method across different modules? **Too few replies**

4.6 IF CRITERION referenced: does more than one person determine and agree the grade boundaries? **69%**

Is assessment timely and progressive throughout the course?

5.1 Does assessment provide a monitor of student performance throughout the module? 59%
5.2 Is there time to allow students to respond to a poor assessment before the end of the module? 50%
5.- Is assessment timely with regard to:
5.3 — the speed with which the results are available to the students? 66%
5.4 — in relation to other assessed work on the module? 69%
[So students know the results and have had feedback before the next piece of assessed work?]
5.5 — in relation to other assessment on other modules? 41%
[So the totality of the assessment as experienced by the students is reasonably distributed and does not all take place in an unreasonably short period.]

Is feedback provided?

6.1 — On all in course assessments? **80%**
6.2 — On the end of module assessment? **25%**
6.3 — To all students? **86%**
6.4 — As written comments sufficiently detailed to enable the student to identify particular weaknesses? **69%**
6.5 — With omissions as well as errors? **69%**
6.6 How do you know that all students access the feedback provided? **43%**
6.7 Are students performing poorly counseled (on a one-one basis)? **54%**
6.8 Does counseling take into account performance on other modules? **45%**

Are resit/second chance arrangements known to students?

7.1 Are these arrangements written, available to the students and explicit with regard to format and material covered? **80%**
7.2 Is the date/time of any resit exams known to the students at least 3 months before it takes place? **46%**
7.3 Are the Learning Objectives the same? **80%**
7.4 Are resit candidates given effective feedback on their performance in the first sit? **42%**

What are the students views on the quality and usefulness of the assessment?

8.1 Are the students views on the assessment processes known and elicited each year? **70%**
8.2 Is this data obtained from all the students other than those absent because of illness? **66%**
8.3 Are they treated as a homogeneous group or are their views fragmented into those of the various sub-groups making up the student body? **62%**
[e.g. year 1 and year 2 taking the same module? Students on different courses but taking the same module?]

THIS ASPECT IS CALCULATED BUT NOT SCORED

Proportion of total teaching time allocated to assessment.

Hours spent by teachers (including demonstrators PG tutors etc) on assessment (hrs)
Hours direct teacher contact with students (hrs)
Hours spent by student being assessed (hrs)
Total each student involved in teaching/learning (lect + prac + selfdirected + directed + tuts others etc) (hrs)
How efficient is the assessment process? *[For each assessed item consider the proportionality between the fraction of the total marks awarded and the assessment time devoted to it. This item is not represented numerically but you should note areas where the time spent by staff in completing the assessment is very large compared with the proportion of marks given for the work]*