The University of Adelaide
School of Dentistry

BDS Assessment Strategy

Background

The content and format of assessment procedures are an integral part of any curriculum and will strongly influence the study behaviour and learning outcomes of the students. It is vitally important that the assessments in the curriculum support our Vision Statements and reflect the core knowledge, skills and behaviours expressed in the Outcome Objectives. They must also provide appropriate motivation and clear direction for the students. This paper advocates the construction of an assessment strategy applied to all years and components of the curriculum using approaches and evidence-based methods reflecting international best practice.

A snapshot survey of student assessment in undergraduate medical education in the UK in 1998 revealed that assessment methods and practices left much to be desired. (Fowell et al, 2000). In particular there appeared to be a lack of knowledge of technical aspects of assessment and limited application of quality assurance procedures. For example, methods of establishing content validity and standard setting were poorly developed. It is likely that a survey of dental schools would show similar results. In addition to meeting expected technical benchmarks there are many other aspects an assessment strategy must address (Van der Vleuten, 1996; Manogue et al, 2002). Such a strategy was developed at the School of Medicine at the University of Sheffield, adopted as policy and successfully implemented within a revised curriculum (Roberts, Newble et al, 2006). It aimed to overcome the previously mentioned technical problems; meet accreditation guidelines set by the General Medical Council and the Quality Assurance Agency; and adopt international best practice methodology (Newble D, Dawson-Saunders B et al, 1994). The latter outlines criteria those responsible for assessment need to achieve to ensure quality procedures. There are 5 steps:

1. Clarification of the purpose of the assessment(s)
2. Defining what is to be tested and determination of the sample size
3. Selecting appropriate test methods
4. Addressing issues of administration and scoring
5. Setting standards for performance

The School of Dentistry has made significant changes to its assessment procedures in recent years. However, with the agreed move to a highly integrated curriculum, further changes will be required. In particular integrated assessments will need to be carefully structured to ensure all content areas are appropriately represented. Meeting the quality standards now expected of assessment procedures, as outlined above, is unlikely to be achieved without a clear strategy and institutional policies. Suggested key features of such a strategy are defined below.

Key Features of the Adelaide School of Dentistry Assessment Strategy

The key features of the strategy are:

- assessments will be closely matched to purpose, program objectives and intended teaching and learning activities
- regular opportunities for formative and self assessment, remediation and counselling will be evident throughout the program
- summative assessments will all be integrated
- assessments of clinical competence (including professional behaviours) will be included in all phases of the program
- the concept of progressive testing will be incorporated into all end-of-course summative assessments (i.e. a proportion of material from previous phases of the program will be included in all subsequent assessments)
- assessment methods will be limited to those on an approved evidence-based list
- assessments will conform to a standardised formats to ensure familiarity to students, to aid staff training and to simplify test development and administration
- summative assessments will meet expected internationally credible standards of validity and reliability
- students will be fully informed of all assessment procedures at the beginning of each year of the program
- university graduate attributes will be assessed
- assessments will comply with University examination policies and regulations

Organisational Aspects

The overseeing of the ongoing development and implementation of the Assessment Strategy will be the responsibility of the Curriculum Committee's Assessment Subcommittee. This will be chaired by a Director of Assessment and consist of the Phase Coordinators, the Theme Directors, an educationist with expertise in assessment, a senior School administrative officer and at least one student representative.

Theme Directors will be responsible for ensuring: that the content of all assessments relating to their theme matches the defined core curriculum outcome objectives; compliance with the policy of progressive testing; that the format of assessments relating to their theme throughout the program complies with the strategy; support of staff development activities; implementation of appropriate quality assurance procedures.

Year Coordinators will be responsible for ensuring that: assessment procedures in each year of the course comply with the strategy; advice is sought on technical aspects; arrangements are made for appropriate staff development activities; assistance is sought for statistical support.

The School Administrative Officer will be responsible for ensuring that: all documents and test materials comply with the "house style"; documentation of assessment procedures is compiled

and made available to students at the beginning of each year; appropriate administrative and financial support is available to those constructing and conducting assessments.

Students will be responsible for: liasing with students and their representatives; providing feedback on student matters relating to assessments; ensuring that all processes and procedures are transparent and fair.

Formative Assessments

Formative assessments are a vital and inherent part of the learning process for students. They provide individual feedback on progress towards learning outcomes and must be conducted independently of summative assessments. A wide range of such activities will be introduced within all components of the program.

Summative Assessments

All components of assessment which the students have to pass, or have to complete before progression from one part of the program to another may occur, are regarded as summative. In general terms they fall into two types: end-of-program examinations and in-program assessments.

Technical Considerations (Compliance with Criteria)

Clarification of Purpose

Conflicts of purpose are often apparent in assessments. Examples include: using assessments primarily intended to give feedback also to provide marks used for summative purposes (thus discouraging students from exposing their deficiencies); attempting to use tests designed to assess competence to also provide grades and classification for distinctions and prizes (thus confusing a criterion-referenced with a norm-referenced approach to testing); and using written tests alone to make judgements about clinical competence (thus ensuring the assessment is invalid).

Defining what is to be tested and the sample size

The first task in approaching the design of an assessment is to define what is to be tested. This is primarily an issue of ensuring content validity. Basically this is to establish that the assessment systematically and representatively samples what it is supposed to measure. A failure to achieve a high level of content validity will undermine the reliability (generalisability) of the assessment and potentially produce serious, unpredictable and often unrecognised distortions to student learning and the curriculum. It is essential for all

parts of the course to have clearly defined outcome objectives to provide the framework against which content validity can be judged. The recommended procedure to achieve this is through the preparation of a **test blueprint** which should be developed and made transparent for all summative assessments. As well as being representative of the content, the assessment must contain a sample size of student performance adequate to achieve acceptable levels of reliability, particularly for high stake assessments. While this is relatively easy to achieve for assessments of knowledge it provides a greater technical and logistical challenge for assessments of clinical skills and professional behaviours.

Selecting appropriate test methods

Broadly speaking, assessment of knowledge can usually be achieved with a written test. Assessments of competence will require, in addition, the use of practical and clinically-based methods both during exams and during the course. Assessment of professional behaviours will require multiple observations over an extended period of time in the real world (eg during group activities, in clinics and placements). While the format of test methods is thought to be less critical than in the past (van der Vleuten, 1996) different methods do have varying strengths and weaknesses and informed choices can be made. Part of our assessment strategy is to attempt to make such choices on the basis of evidence and best practice. As a result the number of test methods to be used throughout the course for summative purposes will be limited to an approved list. This list will be reviewed on a regular basis.

Approved Summative Assessment Methods

Written

- Multiple choice items (1 from 4/5)
- Extended matching items (EMQ)
- Structured short answer questions (e.g. “Key Features” items)
- Modified Essay Questions (MEQ)
- Extended written work (e.g. project reports; poster presentations)
- Log books/Portfolios or “records of achievement”

Clinical/Practical

- Multiple station examinations (e.g. OSCE’s)
- Direct observations of performance/behaviour (e.g. observed cases; mini-CEX)
- Structured reports (e.g. placement assessments)
- Oral presentations (e.g. projects)
- Structured oral examinations

(see References for more details)

Multiple-choice items of the 1 from 4/5 type and extended matching items are generally regarded as the best objective type question formats for providing wide sampling of content knowledge. They are both less subject to guessing and cueing than other formats.

Structured short answer questions of the “key features” type and MEQ’s are favoured for assessing aspects of clinical problem solving.

Logbooks are valuable as a way of ensuring key skills are defined and can provide evidence of completion or mastery of these skills.

Portfolios or “records of achievement” are increasingly being introduced for both formative and summative assessment. However, evidence of reliability for summative use of these methods is limited. (Roberts C, Newble D & O’Rourke A, 2002) Their strength lies in producing an accumulative record of achievement, which can be used for reflection (a formative activity) and for assessment. Electronic versions are being developed. They may be the only practical way of providing evidence about some aspects of competence, particularly those relating to professional behaviour and personal development.

Multi-station exams will provide the backbone of the assessment of practical and clinical skills. However, not all aspects of competence can be validly assessed in this way. Additional (albeit less reliable) approaches will have to be added to ensure validity. These may be included as part of the examination or may have to be undertaken during the course with satisfactory performance being a prerequisite for sitting the end of course assessment or completing the course requirements.

Multiple observations of performance during tutorials, clinic sessions and placements are essential for both formative and summative purposes. Data, particularly that which provides information about professional behaviours, will have to be accumulated over an extended period of time if adequate levels of validity and reliability are to be achieved for this important component of clinical competence. This data may be acquired using a variety of methods such as tutor observation of group performances, observation of cases, mini-CEX encounters (direct observation of selected components of clinical performance on real patients), structured placement reports, and clinical incident records from a range of health professionals.

Oral examination use will be kept to a minimum due to their unfavourable measurement characteristics. If used they will be structured to improve comparability of content and to minimize examiner bias and variability of scoring. They should not be used for high stakes decision-making on their own except as one method for discrimination among candidates for the award of distinction and prizes.

Addressing Issues of Administration and Scoring

Examinations are becoming increasingly complex and costly in order to meet modern quality standards. Our strategy of integration and reducing the number of summative assessments will decrease the overall burden though individually each will be more comprehensive and may become a considerable logistical exercise requiring substantial central administrative support. The harmonisation of methods and adoption of house styles for both written and clinical exams will also improve efficiency.

A range of scoring issues must be tackled. Standardised approaches will be used in all situations through the use of computerised scoring systems, structured rating forms (checklist and global) and marking keys for open response formats. Where examiners are used, every effort will be made to reduce inter-examiner variability and bias through strategies such as training, independent scoring and sharing marking tasks across multiple examiners.

Where multiple test formats are used, attention must be given to the possible confounding effects of combining scores. Equating or conversion to standardised scores may be required.

Weightings within and across components will often have to be applied. These should be considered in the context of validity – more important things should be weighted more heavily to avoid trivialisation (this can apply particularly to scoring schedules for OSCE stations).

Evidence of reliability will be required, particularly for end-of-course summative assessments. This will involve the production of statistical indices, details of which are outside the scope of this paper. High stakes (for the student) decisions increasingly have to be defensible against challenge and litigation, and proof of reliability provides the backbone to this defence.

Setting standards

Universities have traditionally used relative (norm-referenced) standard-setting procedures which are reasonably easy to implement. Increasingly dental schools need to use absolute (criterion-referenced) standards which are more difficult to apply. The latter all require the input of a group of examiners to debate and agree on the expected performance level of borderline students on each item or component of the test.
Well developed methods are available for standard setting for OSCE-type examinations (eg Contrasting Groups methods) which have also been adapted for written tests composed of MEQ's. Absolute standard setting procedures (e.g. Angoff, Ebel or Hofstee) are more difficult to apply but will have to be used for written tests using MCQ's or EMQ's (Norcini, 2003).

Progressive Testing

In an outcome-focussed integrated curriculum all content, irrespective of the subject or its place in the program, is expected to be relevant to the defined outcome objectives. As a result, it is expected that knowledge acquired during courses in earlier parts of the program must be carried forward to later parts. The assessment strategy must encourage integration and discourage the common 'learn and forget' approach adopted by many students. An appropriate way of doing this is by **progressive testing.** The strategy is to have a significant proportion of the content and marks within each assessment based on core material from previous sections of the program. As well as "carrying forward" information from one part of the course to another, progressive testing should also incorporate the concept of assessing performance at a higher level. Consequently, as the program progresses, written and clinical assessments should require students to use their accumulating knowledge of science and their clinical experience to solve problems of increasing difficulty and with increasing levels of competence.

The Portfolio or Record of Achievement

Portfolios are becoming an increasingly popular form of assessment in higher and continuing education and can serve many functions, some formative and some summative. Dental students will benefit from this approach but, in addition, may need to become familiar with the ways in which portfolios could become part of their professional lives after graduation as a component of continuing professional development and revalidation. At undergraduate level, an important function of the portfolio will be as a method of collecting and collating information on aspects of competence not easily assessed in examinations. In particular, this applies to the major component of our Outcome Objectives we call "Professional Behaviours" where reliably identifying students demonstrating unsatisfactory performance is of vital importance. The

majority of these behaviours cannot be appropriately assessed within the framework of an examination, though some limited aspects may be testable. In general, assessments of professional behaviours can only be made on the basis of observations of performance in real world situations. The wide variety of ways in which unsatisfactory professional behaviours may become manifest means that reliable assessment is difficult to achieve. In essence it will require opportunities for observation of performance through extended periods of the course by a large number of assessors (dental, allied health professional, patients) so that repeated examples of unsatisfactory behaviours can be recognised and dealt with, both formatively and summatively. The collection and collation of such data will be enhanced with the introduction of an electronic portfolio.

Quality Assurance

The use of external examiners to monitor the quality of high stakes examinations is generally still highly regarded. With the advent of integrated examinations and new methods of assessment the role of external examiners has changed. No longer can they be chosen simply to represent disciplinary expectations. They must have a broad perspective of the expected outcomes of the dental course and also be competent to evaluate the soundness of the assessment processes and quality control procedures, and their fairness.

Internal mechanisms of quality control must also be in place. Details of the assessments, standard setting procedures and decision-making rules should be made fully transparent. Those responsible for assessment will be required to provide a formal report to a responsible committee which includes psychometric data on both validity and reliability.

References

Fowell SL et.al (2000) Student assessment in undergraduate medical education in the United Kingdom 1998. Medical Education, **34**; (Supplement 1) 1-49.

Van der Vleuten CPM (1996) The assessment of professional competence: developments, research and practical implications. Advances in Health Sciences, **1**; 41-67.

Manogue M etal (2000) Evolving methods of assessment. European Journal of Dental Education, **6**; 41-67.

Roberts C, Newble D, Jolly B, Reed M & Hampton K. (2006) Assuring the quality of high-stakes undergraduate assessments of clinical competence. Medical Teacher, **28**; 536-543

Newble D & Dawson B (Eds) (1994) Guidelines for assessing clinical competence. Teaching & Learning in Medicine, **6**; 213-220.

Tomorrows Doctors, General Medical Council, 2002.

QAA Code of Practice for the Assurance of Academic Quality and Standards (www.qaa.ac.uk). Section 4: External Examining.

QAA Code of Practice for the Assurance of Academic Quality and Standards (www.qaa.ac.uk). Section 6: Assessment of students.

Case SM & Swanson DB (2003) Constructing Written Tests for the Basic and Clinical Sciences (3rd Edition). National Board of Medical Examiners, Philadelphia.

Jolly B & Grant J (Editors) (1997) The Good Assessment Guide. Joint Centre for Higher Education, London.

Newble D. (2004) Techniques for measuring clinical competence: objective structured clinical examination. Medical Education, **38**; 199-203.

Norcini J (2003) Setting standards for educational tests. Medical Education, **37**; 464-469.

Roberts C, Newble D & O'Rouke A. (2002) Portfolio-based assessment in medical education: valid but are they reliable. Medical Education, **36**; 899-900.

(Last modified 6/12/07)

.