

Van Rensburg, V. 2012. The difference between norm-referenced and criterion-referenced assessment methods: What does it mean? Unpublished EDU Resource.

The traditional method of measuring student progress in health professional education used to be written examinations that focused on content knowledge. The global shift in the conceptualisation of curriculum construction, teaching, learning and assessment is evident in new forms of assessment methods currently used in health professional education. Examples of such revised curricula are problem-based, case-based, outcomes-based and competence-based curricula. As part of this shift, thinking about assessment methods and the theory supporting assessment have undergone significant change.

The theoretical premise that traditionally informed our thinking about marks and mark allocation is the psychometric approach. The psychometric approach is framed by the quantitative research paradigm. Psychometrically informed mark allocation is referred to in the literature as 'norm-based' or 'norm-referenced' assessment. To put it simply, it is the system of determining student progress by allocating 'marks for facts'. Norm-based assessments yield results that compare or situate a student's performance within the range of marks obtained by of the rest of the class.

In the current educational practice in the health professions, commensurate with new forms of curriculum construction, teaching and learning, the need arose to assess student learning differently. The newer tendency is to determine student performance by means of the examiner's professional judgement of competence or behavioural-stated outcomes. Knowledge, skills and attitudes are assessed together in an integrated assessment. For example, in an exam of a medical student's proficiency in patient management, the student will be assessed on knowledge, clinical and procedural skills and clinical reasoning but also on communication skills and professional or ethical behaviour. Instead of giving 'marks for facts', pre-determined assessment criteria are used by the examiner to make a professional judgement of the student's proficiency or competence. This type of examination is referred to as 'criterion-referenced' assessment. A criterion-referenced system places less emphasis on how the student performs compared with peers and more emphasis on whether the student can demonstrate and provide evidence of the required level of competence.

Criterion-referenced assessments can be graded by means of scoring rubrics. A scoring rubric, presented in the form of a table, plots assessment criteria against descriptors of competence. Some scoring rubrics do not allocate marks as percentages. Levels of attainment of outcomes or competence are indicated by a limited number of gradations expressed in such terms as 'adequate', 'good', and 'excellent' against each criterion.

Criterion-referenced assessments are underpinned by theories from the social sciences, framed by the constructs of the qualitative research paradigm. Instead of allocating 'marks for facts', typical of the psychometric approach, the examiner makes a professional judgement based on a global impression of evidence of learning by using pre-determined assessment criteria. From the further development of distinguishing between the purpose of assessment, such as whether it is a formative or a summative assessment, a range of 'alternative' or innovative methods of assessment arose. Self-assessment, peer-evaluation, assessment of group work tasks or evidence of reflection on learning are some examples. Other examples are producing a portfolio of learning, posters or health educational pamphlets. 'Marking' of such assessment methods is underpinned by constructs from the

social sciences. For instance, a portfolio of evidence of learning may contain entries of personal reflections on learning experiences. As a reflective entry is a highly personal construction of understanding and meaning making, it is inherently subjective. The subjective nature of such assessment methods poses problems of how it should be assessed and whether validity and reliability are compromised. In the qualitative paradigm, it is argued that subjectivity and personal interpretation can be scrutinised for trustworthiness and credibility to enhance rigour.

To illustrate the differences between norm-referenced and criterion-referenced approaches, they are at times presented theoretically as polar opposites. However, in reality, both norm-referenced and criterion-referenced assessment methods are used in health professional assessments. The choice of method depends on the disciplinary nature, the purpose (formative or summative) and the content covered in the assessment.

It is important to note that criterion-referenced assessments have not replaced norm-referenced assessments, but have added to the repertoire of assessment methods available. The main guiding principle in selecting an assessment method is the concept of 'fitness-for-purpose' or validity of the assessment. Does the assessment method ensure that it measures what it intends to measure? An illustrative example of the need to retain norm-referenced, psychometrically informed assessment methods is the wide use of multiple-choice questions (MCQs) in health professional assessment. MCQs are selected for such purposes as ensuring that a wide range of knowledge is assessed, for reasons of feasibility when assessing large numbers of students, for its high reliability, or for the disciplinary nature of the course being assessed, as for instance in the pre-clinical sciences. No assessment method is perfectly valid and perfectly reliable. Each method has strength and weaknesses. A balanced selection of norm-referenced and criterion-referenced assessment methods will yield a thorough, fair and well-informed impression of a student's capabilities.