Medical Teacher, Vol. 27, No. 4, 2005, pp. 294–297

COMMENTARY

The place of the oral examination in today's assessment systems

MARGERY H. DAVIS[1] & INDIKA KARUNATHILAKE[2]
[1]Centre for Medical Education, University of Dundee, UK; [2]Faculty of Medicine, University of Colombo, Sri Lanka

The oral examination or viva is a traditional form of assessment in which one or more examiners fire questions at the candidate. It typically takes the form of an interview or discussion between the examiners and candidate and happens in an examination hall or other such setting away from patients. It should be distinguished from other types of oral examination such as the long and short case, which take place in the presence of the patient or are focused around a patient seen by the candidate and the oral that is used for defence of written work such as a thesis. The oral examination is said to assess knowledge, to probe depth of knowledge and to test other qualities such as mental agility.

The use of oral examinations in high-stakes assessment systems has been criticized for many years because of low reliability (Colton & Peterson, 1967; Foster *et al.*, 1969; Kelly *et al.*, 1971). The low reliability relates, in part, to the examiner's active participation in the examination, which can introduce bias. In the traditional oral, each candidate may receive a different assessment with regard to content areas addressed, the difficulty of the questions asked, the level of prompting or help provided and the learning outcomes assessed; for example, knowledge of the basic sciences, patient investigation and management. These differences present difficulties not only in a norm-referenced system of assessment, where the intention is to rank the candidates, but also in a criterion-referenced system, where the intention is to assess whether or not the candidate has achieved a pre-determined standard.

The reasons for low reliability also have an adverse impact on validity (Schuwirth & Van der Vleuten, 1996) because of the potential for variation in content matter addressed and in the emphasis given to different content areas.

Oral examinations are usually employed in an attempt to assess the candidate's knowledge of a subject. What may be measured, however, are aspects of a candidate's personality (Bull, 1959). Holloway *et al.* (1967), Holloway *et al.* (1968) and Thomas *et al.* (1992) showed that viva marks correlated with personality scores. Rowland-Morin *et al.* (1991) and Burchard *et al.* (1995) showed that verbal style and dress of the candidates influence oral examination scores. Roberts *et al.* (2000) carried out a discourse analysis (a detailed study of language in use) of the oral component in the membership examination of the Royal College of General Practitioners (MRCGP) and pointed out that candidates from ethnic minorities and those trained abroad may experience particular hidden difficulties with oral examinations leading to discrimination. Furthermore, the discrimination may not be limited to ethnicity. Esmail & May (2000) suggested that candidates from working-class backgrounds and, in some instances, female candidates may also be discriminated against.

The problems with oral examinations extend beyond poor reliability and validity. McGuire (1966) questioned the cost effectiveness of oral examinations, when the cost, in terms of professional time and energy, is weighed against its reliability and validity as a measure of professional competence. Any well-planned examination, however, is costly in terms of examiners' time and effort. The challenge is finding assessment instruments where the effort spent is educationally 'profitable'.

Orals can be highly threatening for candidates with resultant poor performance (Pokorny & Frazier, 1966; Cox, 1982; Thomas *et al.*, 1992; Jolly & Grant, 1997). It can be argued, however, that all examinations are stressful. The question is whether the viva is more stress provoking than other assessments. There is no evidence that orals are more stressful than other exams and, indeed, there is anecdotal evidence to the contrary. Schiff (2001), in a personal narrative, reported that the short case was more stressful than other parts of the MRCP clinical examination.

Norman (2000) pointed out that ''most of the US boards abandoned the oral exam altogether about 30 years ago, based on evidence that, despite intuition to the contrary, it was adding little value to the evaluation process''. Oral examinations, however, continue to be used in specialty board examinations in the UK, Canada and most parts of the former British Empire (Norman, 2000). Why is this?

First, the oral examination is a traditional form of assessment that has been used in undergraduate and postgraduate medical education for many years and breaking with tradition can be difficult. As Jayawickramarajah (1985) suggested, there are ''difficulties in persuading examining boards and training medical examiners to employ appropriate alternative methods''.

Correspondence: Prof. M.H. Davis, Centre for Medical Education, University of Dundee, Tay Park House, 484, Perth Road, Dundee DD2 1LR, UK. Email: m.h.davis@dundee.ac.uk

ISSN 0142–159X print/ISSN 1466–187X online/00/000294-4 © 2005 Taylor & Francis Group Ltd
DOI: 10.1080/01421590500126437