Oral Assessments in Postgraduate Exit Examinations in Forensic Medicine: a Point of View from Sri Lanka

Perera J1, Hulathduwa SR2, Ariyarathne HTDW2

Abstract

Postgraduate education, training and assessment in Forensic Medicine in Sri Lanka are managed by the Postgraduate Institute of Medicine (PGIM), University of Colombo. There are several components in the postgraduate examinations; theory, clinical case presentation, a full autopsy and presentation, Objective Structured Practical Examination (OSPE), Histopathology component and a viva voce examination where a pass grade is compulsory in each component. This includes the 20 minute viva voce examination as well. The viva voce can be used to test intellectual processes such as decision making and justification of opinions; abilities paramount to an effective expert medial witness. However, it has been dropped by many Examination Boards due to several shortcomings. Poor inter examiner reliability is one example that has been quoted as a serious drawback. Certain measures can be taken to increase the validity and reliability of the viva voce, some of which have been introduced in Sri Lanka by the Board of Study in Forensic Medicine, PGIM. Improvements to the viva voce should be an ongoing process. This paper will 1) discuss the value of the viva voce in postgraduate examinations in Forensic Medicine, 2) study its pass/fail reliability as a compulsory component, 3) examine how its shortcomings could be overcome, 4) examine the evolution of the viva voce in postgraduate examinations in Forensic Medicine in Sri Lanka and 5) make recommendations for the future.

Key words: viva voce, examiner reliability

Introduction

Postgraduate education, training and assessment in Forensic Medicine in Sri Lanka are managed by the Postgraduate Institute of Medicine (PGIM), University of Colombo. A separate Board of study in Forensic Medicine facilitates this process (Salgado, 1986). Candidates are selected with a screening test and after two years in-service training in Forensic Medicine and Pathology sit for a diploma examination in Forensic Medicine. Successful trainees sit for the MD in Forensic Medicine following further two year training.

On completion of the MD examination, candidates undergo a two year post MD training in Forensic Medicine and qualify to become PGIM Board Certified Specialists in Forensic Medicine. PGIM certified specialists practice both Forensic Pathology and Clinical Forensic Medicine and work independently at general or base hospitals in Sri Lanka.

In both the diploma and the MD examinations there are several components; theory, clinical case presentation, a full autopsy and presentation, Objective Structured Practical Examination (OSPE), histopathology component and a viva voce examination. Candidates are required to obtain a pass grade in each of the above components to successfully complete the examination. This includes the viva voce examination with a duration of 20 minutes.

The purpose of this paper is to 1) discuss the value of the viva voce in postgraduate examinations in Forensic Medicine, 2) study its pass/fail reliability as a compulsory component, 3) examine how its shortcomings
could be overcome, 4) examine the evolution of the viva voce in postgraduate examinations in Forensic Medicine in Sri Lanka and 5) to make recommendations for the future.

1) The value of the viva voce in postgraduate examinations in Forensic Medicine

Is the viva voce of any value at all?
The viva voce or the oral examination in undergraduate and postgraduate assessments is used by examiners to interact with candidates and assess a variety of intellectual processes in knowledge and decision making (Thomas et al., 1992; Wass et al., 2003). In spite of many criticisms, it is still used in both undergraduate and postgraduate examinations due to its flexibility, apparent fidelity, and its potential for testing higher order cognitive skills (Wakeford et al., 1995). Some institutions, especially the Royal Colleges of the UK, have chosen to retain oral examinations as they feel these add a unique, important and valid dimension to their assessment procedures (Wass et al., 2003).

During oral examinations it is possible to explore understanding by asking particular questions in depth (Cox, 1978). Waas et al. (2003) demonstrated that provided an adequate length of testing time is given, and sufficient independent judgments are made on a wide range of topics, orals can be made psychometrically acceptable.

Structured oral examinations (SOE) were found to be superior to traditional oral exams (Wass et al., 2003) when considering the overall reliability and the inter-rater reliability as revealed in a study of a SOE administered to surgical residents (Anastakis et al., 1991). In the same study, by correlating SOE scores with MCQ and OSCE scores, Anastakis et al., (1991) suggested that the SOE is useful in the assessment of clinical knowledge and problem solving which is all important ability of a forensic medical practitioner. Specialists in Forensic Medicine in Sri Lanka need to communicate decisions/conclusions during their practice to victims of crime and to bereaved relatives. They need to exchange information on a one to one basis with police and other professionals in the criminal justice process. Even more importantly when giving evidence as an expert witness, to perform satisfactorily, a forensic medical practitioner’s evidential and communication skills must be of a high order (Wells & Cordner, 1998). Oral examinations in case presentations or cross examination in moot court situations are currently methods of assessment of these skills at forensic medicine postgraduate examinations in Sri Lanka as well as other parts of the world (Wells & Cordner, 1998).

In Sri Lanka, the areas tested include but are not limited to the candidate’s recognition of the problem, decision making regarding the management, opinion forming ability and the ability to justify and defend them (Jayawickramarajah, 1985). Alternative explanations for the problem/case scenario are propounded to the candidate simulating a cross examination by the defense. The candidate has to consider/exclude giving reasons the possibilities suggested therein. The ability of the candidate to justify these decisions is also tested. The viva voce is one place where flexibility allows exploration of these areas.

Shortcomings of the viva voce
Oral examinations have, however, been criticized for a variety of reasons. The disproportionately high degree of anxiety it produces is higher than in a conventional selection interview (Arndt et al., 1986). A major criticism against the viva voce examination by many authors is its questionable content validity (Thomas et al., 1992; Jayawickramarajah, 1985). In as little as 20 minutes, only a small part of the total knowledge in the field will be subjected to intense scrutiny (Thomas et al., 1992). Another criticism is its low reliability (Jayawickramarajah, 1985) which could be explained by the small sample of examiners and the small samples of content or cases (Delmans, 2001). Low reliability is also in part, related to the examiner’s active participation in the examination, which can introduce bias (Davis & Karunathilake, 2005). At times candidates can be confused by the barrage of questions posed by the examiners or even the statements made to clarify the questions asked, especially when the examiners are not trained in oral examinations.

The disturbing fact is that at times, in oral examinations, candidate attributes that are unconnected to their knowledge or skill can be the determining factor for the grade received. Examiner variability in allocating scores was studied by Thomas et al. (1992) where the video recordings of the oral examination was compared with the transcripts of the same examination, and matched with the non-verbal behaviour of each candidate. Great variability was observed between experienced
examiners in the awarding of marks. It was reported that the students’ level of confidence was a key factor in influencing the oral score. Interestingly in the same study low scores were obtained by men with high levels of anxiety but not by women with similar high levels of anxiety. Therefore, it was hypothesized that examiners may be less tolerant of anxious men than of anxious women (Thomas et al., 1992).

The structured oral component in the MRCGP examination has undergone an extensive process of development with special attention to examiner selection, training and monitoring (Wakeford, 1995). Even here, some features of candidate performance identified by the examiners related more to personal attributes of the candidate rather than knowledge or behaviours acquired in training. Few examples were fluency and creativity (Yaphe & Street, 2003) ‘which fall outside the remit to assess candidate’s professional decision making’. Fortunately however, a candidate’s sex and attractiveness does not seem to affect the marks awarded (Thomas et al., 1992).

The examiners’ decision making process in the allocation of grades in oral examination has been studied. It was demonstrated that examiners generally form a ‘first impression’ as soon as the candidate begins to answer questions, and at an early stage in the orals a ‘provisional grade’ is adjudged before deciding on the final grade after further ‘confirmatory questions’. These grading decisions at various stages were sometimes influenced by candidate’s personal qualities (Yaphe & Street, 2003). In addition to these personal attributes unwittingly determining the awarded grade, discrimination without the examiners being aware of it can at times happen to ethnic minorities, and those trained abroad (Roberts et al., 2000). Physical factors like the time of day when the exam is conducted (Turnbull et al., 1996) or the seating position of the examiner may also alter the score attained by the candidate; in one study, assessors who were face to face with the candidate gave a higher score than those who were listening but could not see the same candidate face to face (Seddon & Pedrosa, 1990).

2) Study its pass/fail reliability as a compulsory component

The reliability of oral examinations is seriously threatened by the examiner (Delmans et al., 2001). Wass et al. (2003) report that the pass/fail reliability of one 20 minute oral with a panel of two trained examiners was as low as 0.70 and even with three examiners it was only 0.7616. In oral examinations even under standardized conditions there can be considerable disagreements about pass/fail decisions according to a study (Holloway, 1968). In this study, it appeared that high extraversion had a greater effect on grades at the viva voce than the level of factual knowledge (Holloway, 1968). Therefore, a multitude of extraneous factors unconnected to the student’s knowledge, skill, practice, or problem solving capability may determine the pass/fail decisions. Because of many such deficiencies the traditional oral examination has been abandoned in many US based examinations (Norman, 2005). With the mounting adverse evidence against oral examinations, the feasibility of retaining it as an exit point where a pass grade is mandatory requires thoughtful consideration.

3) Examine how its shortcomings could be overcome

Substantial work is needed to develop the traditional oral examination into a “best practice oral” form appropriate for high stakes examinations (Davis & Karunathilake, 2005).

Examiners: Examiners should be carefully selected and should have effective interpersonal skills, and the ability to act as an effective member of a small team (Wakeford, 1995). Examiners who can reflect on their performance in grading candidates are more suitable so that they will be aware of their possible biases and will not allow these to interfere with grading decisions. In the study by Yaphe and Street (2003) examiners accepted their personal biases when interpersonal process recall was employed while they studied a videotape of their oral examination. This method can be utilized to identify potential biases of examiners so that biases maybe corrected in order to reduce inter examiner variation (Yaphe & Street, 2003). Ideally their performance should be regularly monitored and feedback given (Turnbull et al., 1996).

Examination process: The oral examination should be structured, have a well-defined goal (Davis & Karunathilake, 2005), examiners’ tasks clearly laid out (Wakeford, 1995) and structured marking schedules employed to increase reliability (Thomas et al., 1992). When considering the appropriate scoring method, a study revealed that global judgments were more reliable in rating the
performance rather than averaged item scores (Delmans et al., 2001). When considering the time duration of the exam, Norman (2000) reports that using more cases with a single examinee per case, and basing the final evaluation on independent assessments by multiple examinees will increase its reliability, a finding endorsed by other authors as well (Wass et al., 2003; Delmans et al., 2001). If a pass/fail decision is still a must in order to successfully complete the examinations, then the viva time should be increased from 20 minutes to at least 40 minutes. This can be done by having 2 orals, 20 minutes each and each exam with a different panel (pair) of examinees (Wass et al., 2003). This will increase the pass/ fail reliability from its present state. Reliability can be further increased (to 0.89) by having three 20 minute vivas with 03 different panels (Wass et al., 2003). However, this is not a feasible option at the final exit examination due to practical time difficulties, but can be explored as an option in the in-service training assessments.

4) Examine the evolution of the viva voce in postgraduate examinations in Forensic Medicine in Sri Lanka

Two decades ago in the postgraduate assessments in Forensic Medicine in Sri Lanka, candidates were tested using a traditional viva voce examination. Each candidate received 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 (Davis & Karunathilake, 2005). Over the years improvements were made with a view to standardize the exam. This included projecting a designated number of multimedia slides on case scenarios which were used as the base for questioning. This was complemented by the introduction of a mock oral examination for the small number of candidates a few years ago by one of the authors thus increasing the candidates’ preparedness. Easy material was used as the first slide (Jayawickramarajah, 1985) to enable the candidate to get acclimatized to the environment. However having projected the slides, in addition to practical questions, factual recall was also tested; aspects which could easily have been tested by MCQ or SAQ, thereby a lost opportunity to test candidates at a higher cognitive level. Standardization was lacking as depending on the given answers, examiners used to question widely around the topic of the slide. Thus different candidates got different content areas in spite of testing with the same slide. In 2012, the following steps were taken to standardize the oral examination.

1. Hawk and dove effect was borne in mind when selecting examiners (McManus et al., 2006).
2. Around seven slides from forensic case scenarios were used and all candidates were tested on the same number of slides.
3. The questions to be asked were discussed and agreed upon by all examiners, written down, and examiners were asked to mark the candidates independently.
4. 1-2 slides or cases were allocated to each examiner. The written questions were given to each examiner.
5. The candidates were made to relax by verbal and non-verbal interaction prior to commencing the exam. (Jayawickramarajah, 1985)
6. Each candidate was asked the same questions in order to improve reliability (Davis & Karunathilake, 2005).
7. The average of the scores was taken as the final score (McManus et al., 2006).

Examiners responded positively and further work is needed to increase the validity and reliability of the viva voce examination.

5) Recommendations for the future

1. Oral examinations in a structured format and with a well defined goal (Davis & Karunathilake, 2005) can be used in postgraduate examinations in Forensic Medicine.
2. However, its place should not be given over to test recall of factual knowledge (Jayawickramarajah, 1985).
3. Seating positions of examiners and candidates should be arranged in a cooperative manner rather than a confrontational manner (Jayawickramarajah, 1985).
4. A pass grade should not be mandatory in the viva voce examination.
5. If a pass is a must at the viva voce then the duration must be increased to 40 minutes. Two 20 minute viva voce could be held with different panels, with at least two examiners each, and each panel testing on different aspects in the curriculum.
6. Proper selection, training and monitoring of examiners should be in place.

7. Feedback should be provided to examiners and further training provided when needed.

References


