

Education

The oral examination in anaesthetic resident evaluation

Chris J. Eagle MD FRCPC,*
Raymond Martineau MD FRCPC,†
Keith Hamilton MD FRCPC‡

Oral examinations have a traditional place in training and evaluation of professionals. Despite a lack of evidence identifying their value in assessment of candidates seeking anaesthetic specialty certification, oral examinations continue to be widely used. Although there is a considerable body of literature concerning oral examinations, there is no description of how this technique is employed in anaesthesia in Canada. The objective of this review is to provide faculty and residents with information concerning the strengths and weaknesses of this format, and the structure of the oral examination as practised in anaesthesia. Reliability of oral examinations can be affected by a number of factors dependent on the examiner, candidate, and the format. Properly constructed and prepared questions have well-defined characteristics. Components tested during the oral examination include: evaluation of a clinical situation, choice of therapy, medical knowledge, ability to deal with emergency situations, decision-making ability, and communication skills. When appropriately planned, the oral examination can be a useful component of the certification process.

Traditionnellement, on utilise les examens oraux pour la formation et l'évaluation des professionnels. Malgré l'absence de validation de leur efficacité, on continue de les soumettre aux candidats qui se présentent aux examens de la spécialité d'anesthésie. Bien qu'on possède une bibliographie importante sur les examens oraux, il n'existe pas de règles formelles sur leur emploi en anesthésie au Canada. L'objectif de cette revue

Key words

EDUCATION: continuing, residents.

From the Departments of Anaesthesia of the University of Calgary,* University of Ottawa,† Dalhousie University,‡ and Chief Examiner,* former,† and current board member‡ of the Oral Examination Board in Anaesthesia (English) of the Royal College of Physicians and Surgeons of Canada.

Address correspondence to: Dr. C.J. Eagle, Department of Anaesthesia, Foothills Hospital, 1403-29 Street N.W., Calgary, Alberta T2N 2T9.

Accepted for publication 18th May, 1993.

consiste à faire connaître aux professeurs et résidents les qualités et défauts de cette formule, et la composition de l'examen oral tel qu'il s'applique à l'anesthésie. La fiabilité de l'examen oral peut être affectée par plusieurs facteurs dont l'examineur, le candidat et le format. Les questions bien préparées et bien exprimées ont des caractéristiques précises. Pendant un examen oral on juge l'évaluation d'une situation clinique, le choix d'une thérapie, la connaissance médicale, l'habileté à réagir en situation d'urgence et à prendre les bonnes décisions, ainsi que la facilité de communiquer. Bien planifié, l'examen oral peut devenir une composante utile au processus de certification.

Oral examinations are the final part of the Royal College certification process in anaesthesia in Canada. The other components are a multiple choice examination (MCQ) and the Final In-training Evaluation Report (FITER). While MCQs are efficient measures of knowledge, they may lack credibility and validity.¹ FITERs assess several components of competence, including clinical and technical skills, problem solving and interpersonal skills, but may lack adequate precision and validity.² Since FITERs are used nationally, there is a potential risk that they may be employed with varying standards among training centres. In the certification process, the oral examination is commonly used to evaluate problem solving ability, communication skills, and knowledge.

Oral examinations have been used in a variety of medical disciplines. Despite the emergence of newer, potentially more valid and reliable, evaluation systems such as the objective structured clinical examination or OSCE (a multistation format),³ oral examinations have persisted. They are a traditional component of the anaesthetic certification process in all of the English speaking world. Perhaps their longevity is owed to the difficulty encountered in developing examination formats which better stimulate the process of anaesthetic care. It is unlikely that oral examinations will disappear in the future despite their well-documented limitations.⁴ It is hoped that this review will aid both candidates and faculty of residency programmes by summarizing the major strengths and

weaknesses of oral examinations, and the structure of the oral examination as commonly used in the certification process in Canada.

Requirements of a measure of clinical competence

When discussing measurement of competence in professionals, three issues should be considered. First, what is meant by the term competence? Second, what measurement tools are available, appropriate, and feasible? Third, what are the strengths and weaknesses of each of these tools? While competence has been defined in many different ways, a practical approach has been to divide competence into five elements.⁵ First, *clinical skills* are needed to obtain relevant information by history and physical examination. Second, *knowledge* is required to interpret these findings and to institute appropriate medical care. Third, *interpersonal skills* are essential to deal effectively with patients and coworkers. Fourth, skills, knowledge, and interpersonal skills must be integrated in order to *solve medical problems*. Fifth, *technical skills* are required to effect patient care or medical procedures. While these five attributes define competence (what the practitioner can do), there is an additional attribute which is important in clinical practice, *performance* (what the practitioner actually does).⁶ To understand the strengths and weaknesses of oral examinations, it is necessary to understand some of the basic terminology of educational assessment.

Validity

Validity "is the degree to which a test 'truly' measures what it is intended to measure."⁷ Validity is the "first priority of any assessment."⁸ There are several ways in which validity may be assessed. First, does the test in question actually cover the area to be tested? That is, is the test comprehensive or does it have *content validity*? For example, an oral examination in anaesthesia which included questions concerning only pulse oximetry would have low content validity. Second, can the test measure the items of interest? That is, does the test have credibility or *face validity*? For example, it would seem unlikely that a multiple choice examination would be able to assess the psychomotor skills required to provide epidural anaesthesia. Third, do higher scores on the test predict success later in the candidate's career? That is, does the test have *predictive validity*? For example, higher scores on a certification examination might be expected to predict the candidates who will become academic, clinical, or professional leaders. Fourth, if there is no gold standard with which to compare the test, does it show results consistent with other expectations? If the test agrees with other credible measures, then it has *construct validity*. For example, do junior residents perform less well than senior residents, or do residents with good in-training assessments perform

better on oral examinations. If the degree of correlation between two measures is very high (0.80–0.90), then they are potentially measuring the same attributes and only one test may be necessary. A correlation of 0.40–0.60 does not by itself indicate that either test is valid, but suggests that both tests may be useful in the overall evaluation system.⁹

Reliability

Reliability is the measure of the "consistency and precision with which a test measures what it is supposed to test."¹⁰ Stated another way, it is a measure of how well the test will measure the same attribute on different days or with different examiners.⁵ There are several methods which can be used to assess reliability. First, tests can be divided into halves and performance in each half compared. This is termed internal consistency. Certification examinations in anaesthesia lend themselves to this type of analysis, since these oral examinations are normally composed of two sections. Reliability between examiners can be evaluated by exercises requiring scoring of videotaped examinations (also a useful training exercise). Videotaped examinations can also be used to evaluate an examiner's consistency over time.

Feasibility

Feasibility is most frequently measured by the cost of the examination. Little information has been published about the relative costs of different examination formats. Certification oral examinations are expensive because of direct costs including examiner travel, accommodation and honoraria, examination preparation, and secretarial support. These are offset to some extent by the fees paid by candidates. There are also unaccounted or indirect costs including the candidates' travel expenses, examiner training, facility preparation, and support costs of the organizing organization. Non-certification examinations also have unaccounted costs, including the training of examiners and mechanisms to relieve both candidates and examiners from clinical duties. While OSCE type examinations have been adopted by some of the examination boards of the RCPSC (Genetics, Obstetrics and Gynecology, and Radiology) apparently in order to improve validity and reliability, it is not clear that these examinations offer any cost advantage. The direct and indirect costs of the RCPSC oral examination are approximately \$2000 per candidate, which is similar to the costs of the OSCEs.* There is some variation in cost based on the number of candidates and the location of examination centre. For comparison, another OSCE, Part 2 of the

*Personal communication: Mrs. Louise Papineau, Examinations Section, RCPSC, Ottawa, May 1993.

Medical Council of Canada Qualifying Examination, has a direct cost of \$1200 per candidate.*

Evaluation of competence in anaesthesia

The certification process in Canada consists of three components which must be satisfactorily completed along with prescribed training periods in specific medical and subspecialty rotations. The first component is a two-session MCQ examination which is composed of approximately 350 questions. The second component is the FITER which is completed in the last months of the candidate's training by the training programme (usually the Residency Director). The FITER requires documentation of the candidate's knowledge, skills, attitudes, and other attributes using rating scales. The training programme is also requested to comment on the candidate's strengths and weaknesses. A satisfactory FITER and written examination score must be achieved before the candidate is allowed to sit the final component, the oral examination, which is two one-hour sessions given on the same day. There are three different examiners present for each session, two principle examiners and one "scribe" who records the questions, responses, and process of the examination. Examiners are assigned responsibility for certain content areas prior to the start of the session. Following each session, the three examiners assign independent grades, followed by discussion of the examinee's performance. A candidate who performs poorly at the oral certification examination of the RCPSC may be subsequently passed if the FITER adequately addresses the weaknesses.

The ABA oral examination consists of two 30-min examinations with two different examiners per session. The questions are predetermined. The four examiners provide independent grades. The Australian and New Zealand College of Anaesthetists (ANZCA) is currently reviewing its examination structure. The previously described system incorporated two oral examinations with two examiners at each session. Each examiner was instructed to provide an independent grade. The two orals contributed 40% of the total points needed to pass the combined written (MCQ and essay) and oral examination.¹¹

Studies of oral examinations

Oral examinations have been reviewed in depth.⁴ Concerns about the validity, reliability, and feasibility of this form of evaluation led the National Board of Medical Examiners¹² and several member boards of the American Board of Medical Specialties (ABMS) to cease this form of evaluation by 1975.⁴ Other boards, including the ABA,

were requested by the ABMS to conduct detailed investigations of the process of their oral examinations. Areas studied by the ABA included examiner performance, assessed by grading standards and the criteria used to designate passing or failing grades, and the relationship between scores on the ABA written and oral examinations. Four examination sessions were studied from 1967-9 and over 1000 candidates were examined. The study found considerable variation in the standard of grading based on the distribution of assignment of failing grades among examiners. Nevertheless, the agreement between examiners evaluating a candidate at a single session was impressive (a reliability coefficient of 0.75). The high reliability achieved in these examinations was felt to be due to a number of factors. First, the evaluation was based on a single examination session. Second, examiner performance was enhanced by a content outline, guidelines, and instructions and training of new examiners. In addition, the performance of examiners was evaluated by periodic observation and discussion of candidates' results. Third, a simple grading system was used. Fourth, the candidates were informed about the examination process in advance.⁴ Many of these features have been incorporated into the oral certification examination in anaesthesia of the RCPSC. The correlation found between the ABA written and oral examinations (0.60) was felt by the authors to indicate that the written and oral measured different aspects of competence. It may be implied from this study that reliable oral examinations can be designed if steps are taken to train examiners and review their performance, control content, and to develop simple consistent scoring systems. The specific items of competence measured by this examination are not known. Although the degree of correlation with written examinations is reassuring, it does not prove that the oral examination is valid, or could not be replaced with more valid or cost-effective measures. However, there is no such examination system in use in anaesthesia.

Reliability of the oral examination may be affected by factors related to the candidate, examiner, or the format. For example, candidates who are excessively nervous may have difficulty answering questions, unrelated to ability. Both the adverse effect of anxiety in candidates and the beneficial effect of examiners in calming candidates have been shown.^{13,14} More recently the effects of verbal and nonverbal communication behaviours have been studied in a blinded multicentre study using videotaped oral examinations in which actors simulated students.¹⁵ When the actors made eye contact and responded at a moderate rate, they were rated higher than when no eye contact was made and a slower response rate was used. Other factors such as fluency have been investigated, but the conclusions are contradictory.^{13,16} Clearly, communication

*Personal Communication: Dr. John Baumber, President, Medical Council of Canada.

skills have an important influence on the perceived level of performance during oral examinations.

Although the effectiveness of specific training programmes for candidates has not been described, most residency programmes have practice orals and some use videotaped review of sessions. Guidelines for candidates are available,¹⁷ as are schema to allow candidates to organize their responses for general and specific problems.^{18,19} An orientation letter is given to candidates by the RCPSC to allow them to understand the structure and overall objectives of the examination.

Training of examiners may produce more uniform delivery of questions and evaluation of performance. In fact, many specialty oral examination boards in both Canada and the United States, including anaesthesia, have adopted formal training programmes in effort to increase both validity and reliability of oral examinations.²⁰⁻²³ The Oral Examination Board in Anaesthesia of the RCPSC has developed an orientation manual, day-long workshops, an examiner evaluation system, and a device to organize content between examination teams and prevent redundancy in the examination (the "grid"). Although some have failed to show improvement after examiner training,²⁴ others have shown that examiner behaviour may be enhanced.²⁰ Examination by pairs of examiners has been found to contribute substantially to the reliability of assigned grades. Checklists have been suggested as a mechanism to reduce the variability in content of questions and grading. It may be that: "The more rigid the structure of oral ... the higher the reliability."²⁵

It is probable that tight control of factors related to the place and procedure of the examinations may also improve the reliability of the examinations. Examinations should commence promptly at the scheduled time and proceed for the allotted time for all candidates. Rooms and waiting areas should be quiet and comfortable in temperature.

In summary, oral examinations have obvious potential weaknesses in comprehensiveness and face validity. There is no published information concerning their predictive validity in anaesthesia. With proper preparation, the tests can be made more reliable and are generally comparable in cost to other measures of similar attributes. It may be that the greatest strength of the oral examination is not as a measurement instrument but as a teaching device.²⁶ There are no studies which demonstrate an improvement in candidates' performance resulting from preparation for oral certification examinations, but candidates and faculty frequently comment on the resulting enhanced command of the specialty.

Components evaluated during the oral examination

The oral examination used in Canadian anaesthetic res-

idency programmes and certification processes is not usually based on check lists of items; rather performance is judged on discussion of a selection of common, verbally presented, clinical situations. These scenarios are chosen to represent the breadth of anaesthetic practice, including both elective and emergency situations, which are used to demonstrate a number of skills (Table). Components evaluated during the oral examination include: evaluation of a clinical situation, choice of therapy, medical knowledge, ability to deal with emergency situations, decision-making ability, and communication skills.

Process of oral examination

Construction of questions

Oral examination questions should be capable of being asked in a few sentences which are clear, unambiguous, uncomplicated, and without repetition. They should not contain "clever" catches or hidden traps. The question should have been thought out clearly beforehand, but not so rigidly that it cannot be changed to suit the candidate's response. This requires each question to have a decision tree prepared. The examiner must have clear expectations of the important points that the candidate should make during the answer. Ideally, the question should be practised on colleagues before being used in examinations. It is not unusual for there to be a much greater diversity in response than is initially recognized by the question's author.

The clinical scenario should be extremely clear, with as few layers as possible. For example, discussion of a 65-year-old man with angina pectoris who is suffering from an acute upper GI bleed and is hypovolaemic will give the candidate ample opportunity to discuss a number of resuscitation, monitoring, and management issues. Addition of a difficult airway, chronic renal failure, or malignant hyperthermia results in a situation so complex that no additional information may be obtained about the candidate. In addition, the possibility of introducing confusion is increased. Useful questions are usually not dependent on reciting lists. For example, the question: "What are the anaesthetic implications of rheumatoid arthritis?" may result in an excessively long response which reveals little about the candidate's ability to make decisions. However, it may be a useful question if the objective is to help a nervous candidate start the examination or to allow a candidate to regain composure.

Questions that are dependent on one key clue or action may be difficult to assess if the candidate does not identify the root issue. For example, using a question based on the surgical procedure of Harrington rod insertion to discuss management of intraoperative pneumothorax would be ineffective if the candidate became sidetracked by other

TABLE Problem solving as tested in the Oral Examination

<i>Area tested</i>	<i>Components</i>	<i>Common deficiencies in candidates</i>
Evaluation of a clinical situation	Collection of appropriate historical, physical examination and laboratory information	Incomplete data collection
Choice of therapy and strategy, including alternatives	Ability to choose a logical plan while displaying knowledge of other available options	Lack of consideration of viable options
Rationale for selection of a given technique	Knowledge of advantages, disadvantages, and complications of a given technique	Deficiencies in knowledge about a chosen technique, or choice of a technique or plan which would result in a dangerous situation
Ability to deal with emergency situations	Rapid, assertive response to immediate care and prioritizing therapy	Reluctance to react to a situation with adequate urgency or failure to consider <i>all</i> of the problems present in the situation
Decision-making ability	Ability to make a rapid assertive response	Management plan developed with incomplete consideration of available data, inability to incorporate new data, or excessive hesitancy in choosing a plan; prioritize appropriately
Communication skills	Ability to provide a lucid, complete response	Continuous prompting required by examiner
Consultant skills	Ability to provide advice to a colleague or patient	Failure to provide adequate information or an appropriate breadth of options

features of the case. Finally, a good question is not esoteric. A scenario based on a pregnant achondroplastic dwarf for intracranial aneurysm clipping and Caesarean-section, might be an interesting exercise but is not useful for examinations.

Types of questions

There are several different types of questions which are useful in different situations. Usually, each session requires an introductory or opener question, designed to be relatively non-stressful, and to help put the candidate more at ease. It is not used to make critical decisions about performance on the examination. Examples include:

*What are the CAS Guidelines for intra-op monitoring?
What are the contra-indications to thiopentone?
What are your NPO guidelines for children?*

Consultant type questions have a clinical scenario, a moderate pace, and room for discussion which may be truncated at the discretion of the examiner. For example:

A severe asthmatic presents for elective cholecystectomy. How would you assess and manage this patient?

These situations can be easily adapted to acute medical problems:

Intraoperatively, this patient has now developed increased inspiratory pressures and unilateral air entry. The O₂ saturation is 89%. What would you do?

Action questions require a set of decisions and actions with little discussion, the emphasis is on working quickly through a clinical scenario:

You are called STAT to the case room to manage a major antepartum haemorrhage, shocky mother, and flat fetal heart tracing.

You are called to provide a cardioversion in CCU for a man who had a myocardial infarction two days ago, and now has developed ventricular tachycardia one hour after lunch.

The quick snapper questions can be useful to assess the candidate who is having difficulty making clinical decisions or taking action. They tend to be brief action questions and are intense and fatiguing for the candidate.

Shortly after you give penicillin to your patient the mean arterial blood pressure is 50 mmHg.

You are called to Recovery Room to see a cyanosed patient who has just had a radical neck dissection.

You walk into the OR as your resident gives 100 mg of succinylcholine to a patient with one-week-old third degree burns to 60% of his body surface area.

Occasionally it is necessary to ask questions which are intended to address problems related to the process of the examination rather than specific content issues. Examples of process control questions include:

"What are your specific concerns as you approach this patient?" to pin down a candidate who is rambling.

"The patient with the chronic subdural whom you have been assessing suddenly loses consciousness and dilates a pupil," to change the time scale for a candidate who is making overly theoretical responses.

Poor performance of the candidate

The examiner will be faced with candidates who perform poorly. In such circumstances, the examiner must try to determine the candidate's comprehension of the question asked. Focusing the question on the most important elements will sometimes allow the examiner to calm the candidate and eliminate any possibility of confusion. In addition, the examiner may slow the pace of the examination for a period to prevent the candidate from feeling intimidated. A framework, based on the categories shown in the Table, may be useful in analyzing performance problems.

Summary

Although concerns about the reliability and validity of oral examinations persist, a number of methods can be employed to increase the effectiveness of this form of examination. These consist of proper preparation of examiners and questions, standardization of grading, and uniformity in the process of the examination. These methods decrease the subjectivity inherent in the oral. Oral examinations are used extensively in anaesthesia in the English speaking world with apparently effective results.

Acknowledgements

The authors would like to acknowledge the encouragement and leadership give by Dr. Frank King, former Chief Examiner of the English Oral Examination Board, over a period of several years and the wisdom of present and past members of the English and French Oral Examination Boards in Anaesthesia of the RCPSC.

References

- 1 Neufeld VR. Written examinations. *In: Neufeld RV, Norman GR (Eds.). Assessing Clinical Competence.* New York: Springer Publishing Co., 1985; 110.
- 2 Streiner DL. Global rating scales. *In: Neufeld RV, Norman GR (Eds.). Assessing Clinical Competence.* New York: Springer Publishing Co., 1985; 138.
- 3 Harden RM, Gleason FA. Assessment of clinical competence using an objective structured clinical examination (OSCE). *Med Educ* 1979; 13: 41-54.
- 4 Muzzin LJ, Hart L. Oral examinations. *In: Neufeld RV, Norman GR (Eds.). Assessing Clinical Competence.* New York: Springer Publishing Co., 1985; 71-93.
- 5 Norman GR. Defining competence: a methodological review. *In: Neufeld RV, Norman GR (Eds.). Assessing Clinical Competence.* New York: Springer Publishing Co., 1985; 15-35.
- 6 Senior JR. Toward the measurement of competence in medicine. Philadelphia: National Board of Medical Examiners, 1976: 15.
- 7 Neufeld VR. An introduction to measurement properties. *In: Neufeld RV, Norman GR (Eds.). Assessing Clinical Competence.* New York: Springer Publishing Co., 1985; 45.
- 8 Newble D. A Handbook for Medical Teachers. 2nd ed. Norwell, MA: MTP Press, 1987; 83.
- 9 Kelley PR, Mathews JH, Schumacher CF. Analysis of the oral examination of the American Board of Anesthesiology. *Journal of Medical Education* 1971; 46: 982-8.
- 10 Newble D. A Handbook for Medical Teachers (2nd ed.). Norwell, MA: MTP Press, 1987; 90.
- 11 Fisk GC, Dunstan MR, Currie TT, Dwyer BE, Newman HC. The final FFARACS examination. *Anaesth Intensive Care* 1975; 3: 227-33.
- 12 Hubbard JP, Levitt EJ, Schumacher CF, Schnabel TG. An objective evaluation of clinical competence. *N Engl J Med* 1963; 272: 1321-8.
- 13 Pokorny AD, Frazier SH Jr. An evaluation of oral examinations. *J Med Educ* 1966; 41: 28-40.
- 14 Waugh D, Moyse CA. Medical education II: Oral examinations: A video study of the reproducibility of grades in pathology. *Can Med Assoc J* 1969; 100: 635-40.
- 15 Rowland-Morin PA, Burchard KW, Garb JL, Coe NP. Influence of effective communication by surgery students on their oral examination scores. *Acad Med* 1991; 66: 169-71.
- 16 Evans L, Ingersoll RW, Smith EJ. The reliability and taxonomic structure of the oral examination. *J Med Educ* 1966; 41: 651.
- 17 Gallagher CJ, Lubarsky DA. Preparing for the Anesthesia Orals: Board Staff. Toronto: Butterworths, 1989.
- 18 Bready LL, Smith RB. Decision Making in Anesthesiology. Toronto: B.C. Decker Inc., 1987.
- 19 Hardy J-F, Guay J, Baillargeon R. A plan for the anaesthetic management of a patient coming to surgery (Letter). *Can J Anaesth* 1986; 33: 256-62.
- 20 DesMarchais JE, Jean P, Delorme P. Training in the art of asking questions at oral examinations. *Annals RCPSC* 1989; 22: 213-6.
- 21 Krome RL, Wagner DK, Munger BS. Standardization of oral examinations. *In: Lloyd JS (Ed.). Oral Examinations in Medical Specialty Board Certification.* Chicago: American Board of Medical Specialties, 1983; 25-53.
- 22 Trier WC. Oral examiner training by the American Board of Plastic Surgery. *In: Lloyd JS (Ed.). Oral Examinations in Medical Specialty Board Certification.* Chicago: American Board of Medical Specialties, 1983; 79-92.

- 23 *Stevens WC*. Training of oral examiners: the oral examination workshop of the American Board of Anesthesiology. *In*: Lloyd JS (Ed.). *Oral Examinations in Medical Specialty Board Certification*. Chicago: American Board of Medical Specialties, 1983; 73–8.
- 24 *Newble DI, Hoare J, Sheldrake PF*. The selection and training of examiners for clinical examinations. *Med Educ* 1985; 19: 290–3.
- 25 *Muzzin LJ, Hart L*. Oral examinations. *In*: Neufeld RV, Norman GR (Eds.). *Assessing Clinical Competence*. New York: Springer Publishing Co., 1985; 80.
- 26 *Abrahamson S*. The oral examination: the case for and the case against. *In*: *Proceedings: Conference on the oral examination*. Des Plaines, Ill: The American Board of Medical Specialties, 1975; 25–7.