

Oral Exams

What works and what doesn't

CESMA

Innsbruck

24 May 2014

Nigel Bax

An Evaluation of Oral Examinations

ALEX D. POKORNY, M.D.* AND SHERVERT H. FRAZIER, JR., M.D.†

Baylor University College of Medicine, Houston, Texas

Oral examinations are in widespread use today in the evaluation and grading of medical students, residents, Ph.D. candidates, and others. They are of particular interest to psychiatrists because a major portion of the certification examinations of the American Board of Psychiatry and Neurology has consisted of oral examinations. Although there is widespread criticism of the reliability, validity, and usefulness of oral examinations, there are few objective studies in the literature concerning their use in medical settings. The present investigation was designed to explore several aspects of the oral examination process and procedure, particularly its validity and the agreement among different examiners.

agreed (it was suggested that it may be "personality"). The correlation between the written and interview examinations was poor. Bull reported that during one final examination, an emeritus professor of logic, who had no medical training, interviewed the students for ten minutes each on nonmedical topics, and gave them grades which correlated +.45 with the final mark in medicine.

Goldstein (2) has criticized the practice of giving class rankings on the basis of examinations. He concluded that ranking grades were fallacious, contributed nothing positive, and fostered undesirable student attitudes. In one experiment Goldstein showed that one-third of a class of students was subject to a rank shift amounting to 2 full quartiles, de-

An Evaluation of Oral Examinations

ALEX D. POKORNY, M.D.* AND SHERVERT H. FRAZIER, JR., M.D.†

Baylor University College of Medicine, Houston, Texas

Oral examinations are in widespread use today in the evaluation and grading of medical students, residents, Ph.D. candidates, and others. They are of particular interest to psychiatrists because a major portion of the certification examinations of the American Board of Psychiatry and Neurology has consisted of oral examinations. Although there is widespread criticism of the reliability, validity, and usefulness of oral examinations, there are few objective studies in the literature concerning their use in medical settings. The present investigation was designed to explore several aspects of the oral examination process and procedure, particularly its validity and the agreement among different examiners.

agreed (it was suggested that it may be "personality"). The correlation between the written and interview examinations was poor. Bull reported that during one final examination, an emeritus professor of logic, who had no medical training, interviewed the students for ten minutes each on nonmedical topics, and gave them grades which correlated +.45 with the final mark in medicine.

Goldstein (2) has criticized the practice of giving class rankings on the basis of examinations. He concluded that ranking grades were fallacious, contributed nothing positive, and fostered undesirable student attitudes. In one experiment Goldstein showed that one-third of a class of students was subject to a rank shift amounting to 2 full quartiles, de-

‘The correlation between the written and interview examinations was poor.’

Pokorny and Frazier, 1966

AN EXAMINATION OF EXAMINATIONS

THE pamphlet recently issued with this title by Sir Philip Hartog and Dr. E. C. Rhodes,¹ on behalf of the International Institute Examinations Enquiry, has received considerable publicity in the Press and has aroused in the general public a good deal of interest, and in many quarters disquietude and indeed alarm. For the benefit of those who have not read the pamphlet itself it may be well to explain what body it is that has undertaken the investigations summarised in this report. In May 1931 there assembled at Eastbourne an international conference on examinations under the auspices of the Carnegie Corporation, the Carnegie Foundation, and the International Institute of Teachers College, Columbia University, U.S.A. Representatives of England, Scotland, France, Germany, Switzerland, and the U.S.A. were present at the meeting. As a result of the conference committees were set up in each of the afore-mentioned countries. They worked independently, and presented their

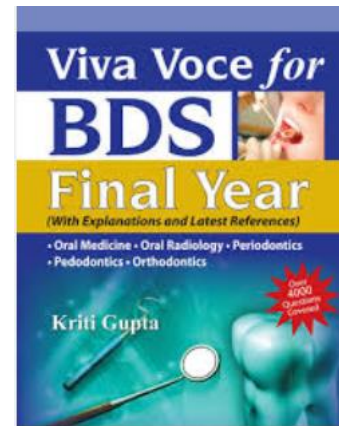
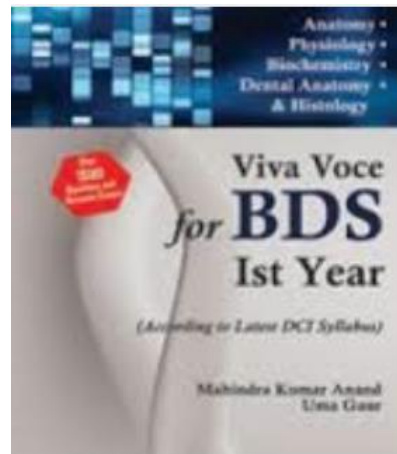
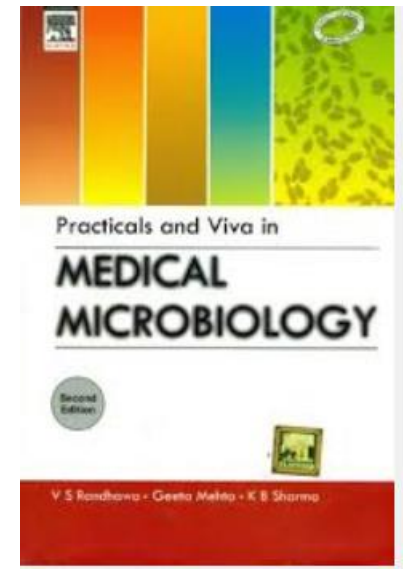
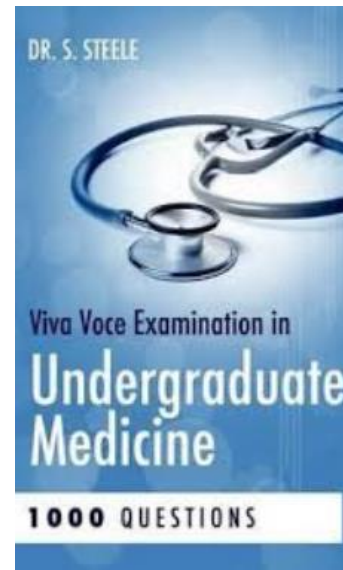
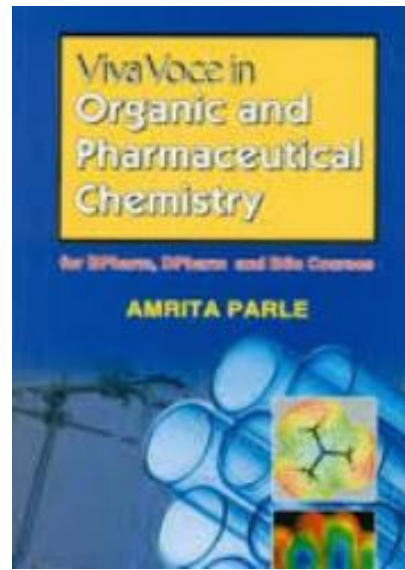
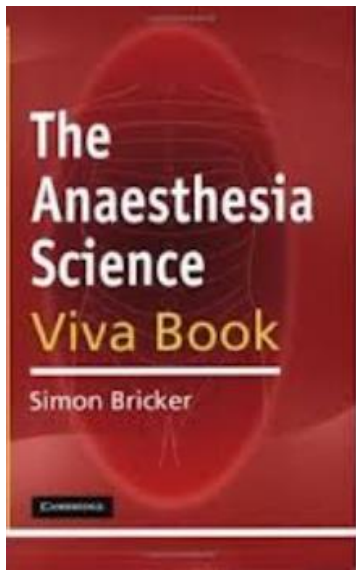
AN EXAMINATION OF EXAMINATIONS

15 scripts with 15 expert markers

Low concordance between 1st and 2nd marking

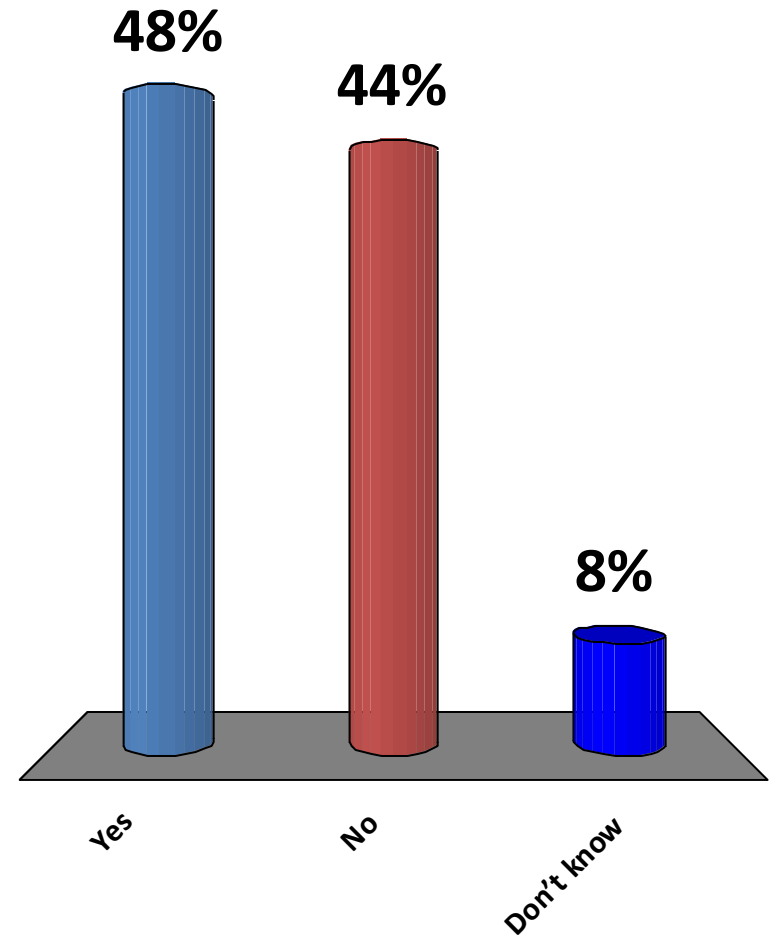
Examiners change their minds often

Pass/fail boundary not the same with all markers



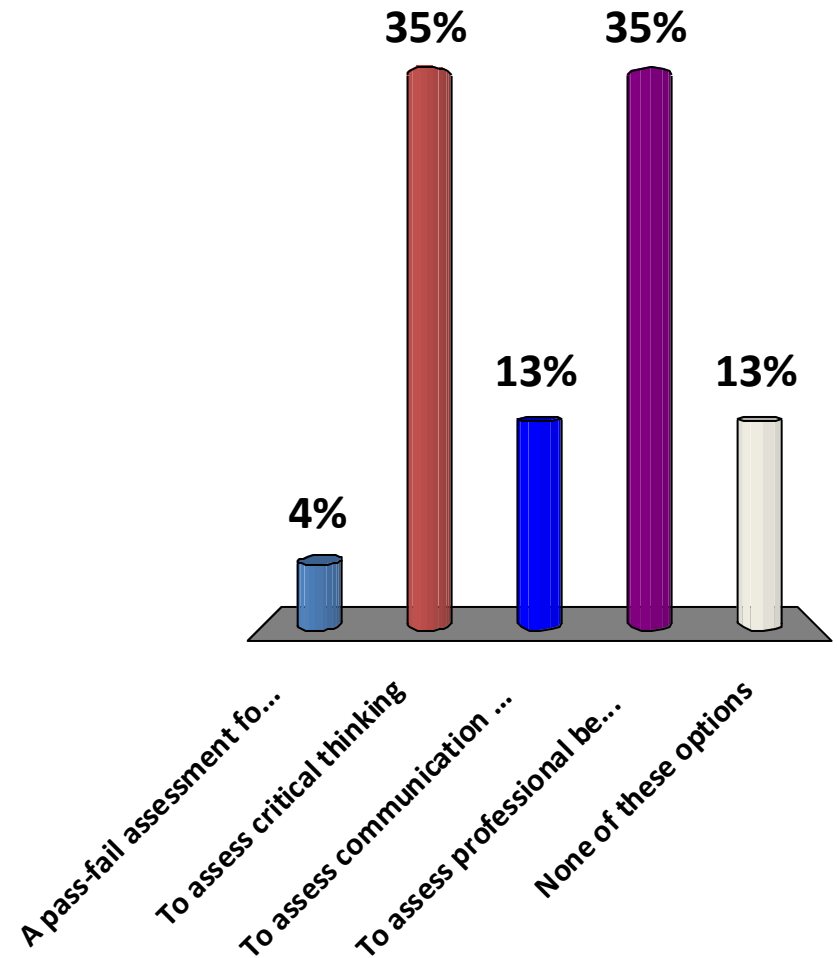
An oral examination must be a part of all UEMS assessments.

- A. Yes
- B. No
- C. Don't know



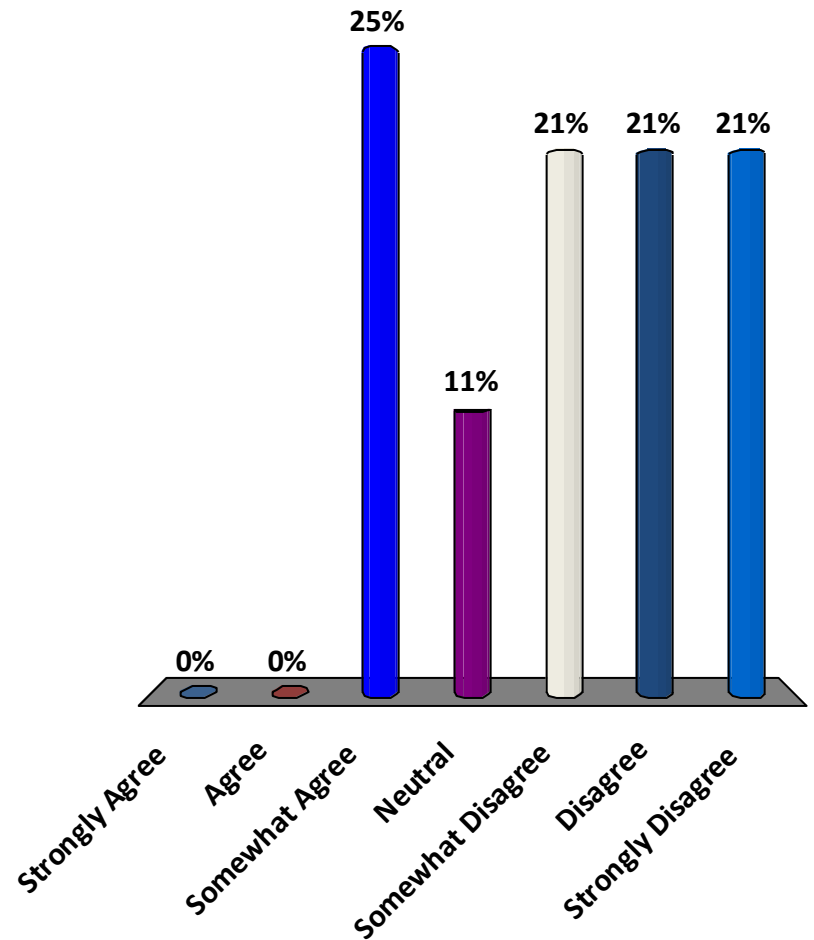
What is the main purpose of an oral examination?

- A. A pass-fail assessment for borderline candidates
- B. To assess critical thinking
- C. To assess communication skills
- D. To assess professional behaviours
- E. None of these options



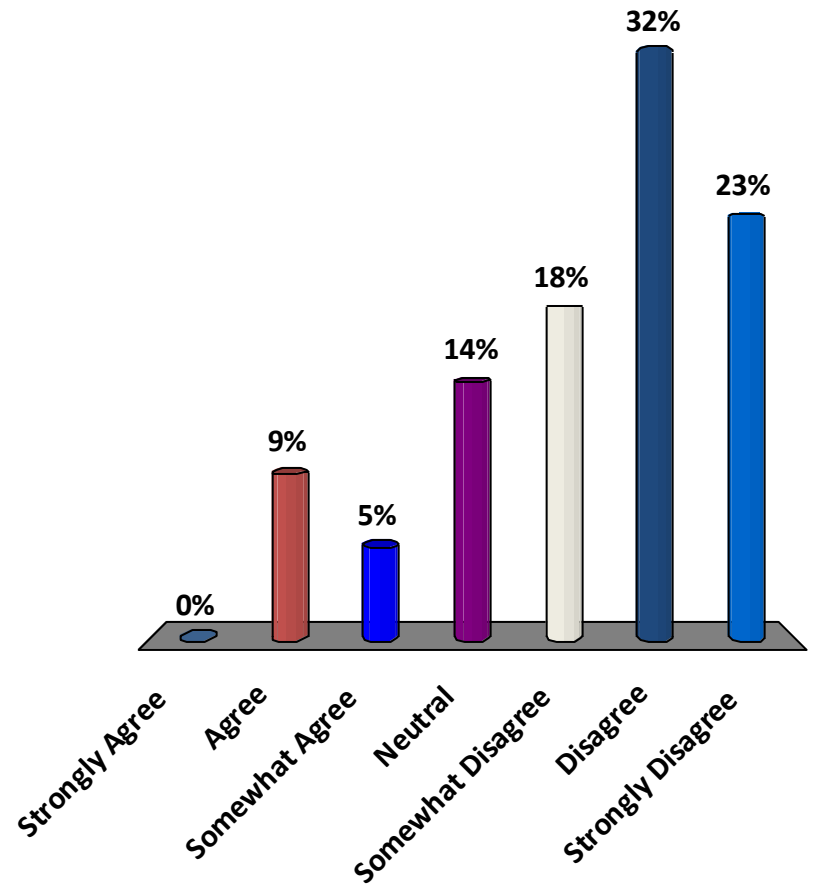
An oral examination has high statistical reliability.

- A. Strongly Agree
- B. Agree
- C. Somewhat Agree
- D. Neutral
- E. Somewhat Disagree
- F. Disagree
- G. Strongly Disagree



An oral examination has equal statistical reliability to a 12 station OSCE

- A. Strongly Agree
- B. Agree
- C. Somewhat Agree
- D. Neutral
- E. Somewhat Disagree
- F. Disagree
- G. Strongly Disagree



Reliability of assessments for different testing times

	1h	2h	4h
Multi-choice	0.62	0.76	0.93
Oral examination	0.50	0.69	0.82
Long case	0.60	0.75	0.86
OSCE	0.54	0.69	0.82
Mini-CEX	0.73	0.84	0.92
Clinical practice video	0.62	0.76	0.93

Consistency, Inter-rater Reliability, and Validity of 441 Consecutive Mock Oral Examinations in Anesthesiology

Implications for Use as a Tool for Assessment of Residents

Armin Schubert, M.D.,*† John E. Tetzlaff, M.D.,†‡ Ming Tan, Ph.D.,§ J. Victor Ryckman, M.D.,‡
Edward Mascha, M.S.||

Background: Oral practice examinations (OPEs) are used extensively in many anesthesiology programs for various reasons, including assessment of clinical judgment. Yet oral examinations have been criticized for their subjectivity. The authors studied the reliability, consistency, and validity of their OPE program to determine if it was a useful assessment tool.

Methods: From 1989 through 1993, we prospectively studied 441 OPEs given to 190 residents. The examination format closely approximated that used by the American Board of An-

esthesiology. Pass-fail grade and an overall numerical score were the OPE results of interest. Internal consistency and inter-rater reliability were determined using agreement measures. To assess their validity in describing competence, OPE results were correlated with in-training examination results and faculty evaluations. Furthermore, we analyzed the relationship of OPE with implicit indicators of resident preparation such as length of training.

Results: The internal consistency coefficient for the overall numerical score was 0.82, indicating good correlation among component scores. The interexaminer agreement was 0.68, indicating moderate or good agreement beyond that expected by chance. The actual agreement among examiners on pass-fail was 84%. Correlation of overall numerical score with in-training examination scores and faculty evaluations was moderate ($r = 0.47$ and 0.41 , respectively; $P < 0.01$). OPE

This article is accompanied by an Editorial View. Please see:
James FM III: Oral practice examinations: Are they worth it?
ANESTHESIOLOGY 1999; 91:4-6.

The oral examination: a study of academic and non-academic factors

C. S. THOMAS, G. MELLOP, K. CALLENDER, J. CRAWSHAW, P. M. ELLIS, A. HALL, J. MACDONALD, P. SILFVERSKIOLD & S. ROMANS-CLARKSON§

Department of Psychological Medicine, Wellington School of Medicine, University of Otago and §Department of Psychological Medicine, Dunedin School of Medicine, University of Otago

40 medical students – 6 consultant psychiatrists

Spoken content of viva is not main determinant of outcome

Poor agreement between examiners

Relationship between score and candidate 'confidence'

Does a 30 minute oral examination have content validity?

NO Morley and Snaith, 1989

Is there high inter-examiner agreement?

NO Wilson et al, 1969

When the inter-examiner agreement is high is there agreement over pass/fail decisions?

NO Holloway et al, 1968

Does examinee personality influence the mark?

YES Marshall and Ludbrook, 1972; Thomas et al, 1992

Viva Voce Assessment

Guidance on viva voce examinations as a secondary form of assessment

Principles

These Guidelines are applicable to undergraduate students only, and relate to vivas undertaken as a secondary form of assessment specifically to reflect traditional expectations within a subject area. The University of Sheffield's Code of Practice for [External Examiners](#) of Taught Courses states the following in relation to the role of examiners participating in the award of degree classifications (section 2.4):

"It is University of Sheffield policy that viva voce examinations will not be held, except where they are a primary aspect of the assessment process. Thus, they will not be used to help determine a degree classification where all the marks are already available. Exceptions will be subject to approval by the Appropriate Faculty Officer or the Pro-Vice-Chancellor for Learning and Teaching in the case of collaborative programmes."

The following guidance therefore does not apply to viva voce examinations which are used as a primary means of assessing students' knowledge, for example those used in language and medical degrees.

"It is University of Sheffield policy that viva voce examinations will not be held, except where they are a primary aspect of the assessment process. Thus, they will not be used to help determine a degree classification where all the marks are already available. Exceptions will be subject to approval by the Appropriate Faculty Officer or the Pro-Vice-Chancellor for Learning and Teaching in the case of collaborative programmes."

DANGERS OF ORAL EXAMINATIONS

- Language differences
- Comprehension differences
- Potential for discrimination
- Inherently less reliable than other assessment methods
- Challengeable
- A cheap way of having a surrogate for a clinical examination

BENEFITS OF ORAL EXAMINATIONS

An oral examination must be a part of all UEMS assessments.

- A. Yes
- B. No
- C. Don't know

