REVIEW ARTICLE NORTH BENGAL MEDICAL COLLEGE JOURNAL [2018], VOL., 4, No. 1 : 31-36

Effectiveness of Oral Examination as an Assessment Method in Medical Education

*S M Akram Hossain,[1] S M Mosaddeq Hossain[2]

ARTICLE INFO
Article history:
Received: 19 September 2017
Accepted: 08 November 2017

Online:
www.nbmc.ac.bd

Keywords:
Assessment methods, Medical education, Oral exam

ABSTRACT
An oral Examination is a form of assessment where a set of stimulus questions are developed that address critical areas of knowledge or sets of abilities related to a competency or set of competencies. Students are expected to respond verbally in their own words, which allow an assessment of student's depth of comprehension, and capacity to apply knowledge and insights to different situations. Responses to the questions are assessed using a rating scale or scoring system. In practice, oral exams were used not as a substitute, but as a complement to written exams. They are a way to ask what is not feasible through the written format. Traditional oral examination has been changed to structured form to ensure greater reliability. Even then, teachers are not yet building up to conduct oral exam in such a structured way. Examiners differ in their personality, style and level of experience with variation of questioning and scoring from student to students. Weakness of reliability on oral examination still exists. Students also feel very stressful during the oral examinations. Two concepts are important to appreciate about the methods used for assessment. It must be valid, namely that what is measured truly reflects the trainee's ability. Second, a method should be reliable so that if repeated multiple times with the same resident the results are consistent. Worlds' leading medical schools now-a-days used oral examination only for borderline and distinction students. Substantial work, however, is needed to develop the prevailing oral exam into a 'best practical oral' format appropriate for medical education.

[1.] *Professor, Department of Anatomy, North Bengal Medical College, Sirajganj*
[2.] *Resident, Department of Pathology, Rajshahi Medical College, Rajshahi*

***Corresponding author:** ✉ akhossain_17@yahoo.com*

INTRODUCTION

An assessment is defined as to include all activities that students undertake in the classroom that can be used to modify students' learning. It also include observation of students made by teachers in the classroom discussions, analysis of work done by students in classroom, homework tasks and tests.[1] Competence in medical science is defined as "the habitual and judicious use of communications, knowledge, technical skills, clinical reasoning, emotions, values, and reflection in daily practice for the benefit of the individuals and communities being served".[2,3] Competence is not an achievement but rather a habit of lifelong learning[3] and it is contextual, reflecting the relationship between person's abilities and the tasks he or she required to perform in a particular

situation in the real world.[4] Medical educators around the world have successfully used many methods of assessing learners, both written and oral.[5] The oral method of assessment was defined by Joughin as an "assessment in which a student's response to the assessment task is verbal, in the sense of being expressed or conveyed by speech instead of writing".[6] He also pointed out that oral examinations enable two qualities to be measured: student command of the oral medium, and a student's command of content.[6] Teaching and learning are mutual procedures that influence each other and assessment determines both students and teachers efforts.[7] The purpose of assessment is to meet the public expectation about the quality graduates, to give feedback to the educational stakeholders about the curriculum, to differentiate the students practical ability according to their efficiency and to monitor their own learning. Assessment also provides a high degree of fairness and objectivity in testing and produces data to enable continuous quality improvement.[8] The welfare and indeed the future health of people depend on the quality of medical graduates and the quality of medical graduates depends on effectual medical education. Educationists believe that, simply by changing the assessment style for the learners can affect the way of engagement of students with the subject contents.[9] Assessment drives learning and learning drives practices. This review aimed to highlight the prevailing oral assessment pattern and to suggest further changes in its system in order to produce competent medical practitioners to meet the health needs of the community.

Assessment tools

Assessment tools comprise a wide range of instruments and methodologies designed to gather this information for feedback, diagnostic purposes, and identifying successful attainment of competence. The utility or usefulness of an assessment has been defined as a product of its validity, reliability, cost-effectiveness, acceptability and educational impact.[10]

Validity

Validity "is the degree to which a test 'truly' measures what it is intended to measure". Validity is the "first priority of any assessment".[11] Validity essentially deals with the design aspect of an assessment. An assessment should fulfill the objective for which it is designed. The objective of the assessment is to facilitate a student's capacity to demonstrate the knowledge, skills and they acquire in relation to the subject being assessed.

Types of validity

Content validity

It measures the extent to which the content of the test matches the instructional objectives. For example, if the final exam includes content covered only during last six weeks, it is not a valid measure of the course's overall objectives, that is, it has a very low content validity.[10, 11]

Criterion validity

It determines the extent to which scores on the test are in agreement with (concurrent validity) or predict (predictive validity) an external criterion. For example, if the end-of-year final exams in a university correlate highly with the national competitive exam, they would have concurrent validity.[10, 11]

Construct validity

It determines the extent to which an assessment corresponds to other variables, as predicted by some rationale or theory. If you can correctly hypothesize that English for speakers of other languages (ESOL) students will perform differently on a reading test than English-speaking students (because of theory), the assessment may have construct validity.[10, 11]

Reliability

Reliability relates to consistency in measurement, that is, scores derived from a reliable assessment tool are similar across assessment events.

Reliability is of central importance in assessment because trainees, assessors, regulatory bodies and the public alike want to be reassured that assessments used to ensure that students are competent enough and would reach the same conclusions if it were possible to administer same test again on the same student in the same circumstances. Reliability is typically reported as value ranging from 0.0 to 1.0. Reliabilities above 0.90 are considered to be excellent, Reliabilities below 0.70 are considered suspect, and results from such an assessment tool should be interpreted with caution.[10]

Types of reliability

Inter-rater or inter-observer reliability

It is used to assess the degree to which different raters/observers give consistent estimates of the same phenomenon. It can be calculated by measuring the percent of agreement between the raters, or calculating the correlation between the ratings of the two observers.[13]

Test-retest reliability

It is used to assess the consistency of a measure from one time to another, given same assessment twice, separated by days, weeks, or months. Reliability is stated as the correlation between scores at Time 1 and Time 2.[13]

Parallel-forms reliability

It is used to assess the consistency of the results of the two tests constructed in the same way from the same content domain. To measure, create two forms of the same test (vary the items slightly). Reliability is stated as correlation between the scores Test 1 and Test 2.[13]

Internal consistency (Alpha,a)

It assesses the consistency of the results across items within a test. Compare one half of the test to the other half. Or, use methods such as Kuder-Richardson Formula 20 (KR20) or Cronbach's Alpha. Coefficient alpha and KR-20 both represent the average of all possible split-half estimates. The difference between the two arises when they would be used to assess reliability.[13]

Relationship between reliability and validity

If a test is unreliable, it cannot be valid. For a test to be valid, it must be reliable. However, because a test is reliable does not mean it will be valid. Hence, reliability does not imply validity. In terms of accuracy and precision, reliability is analogous to precision, while validity is analogous to accuracy.[5,6,10,11,13]

Effectiveness of oral Exams

The oral exam format enables instructors to test the students on all five cognitive domains of Bloom's taxonomy (Figure 1).[14]

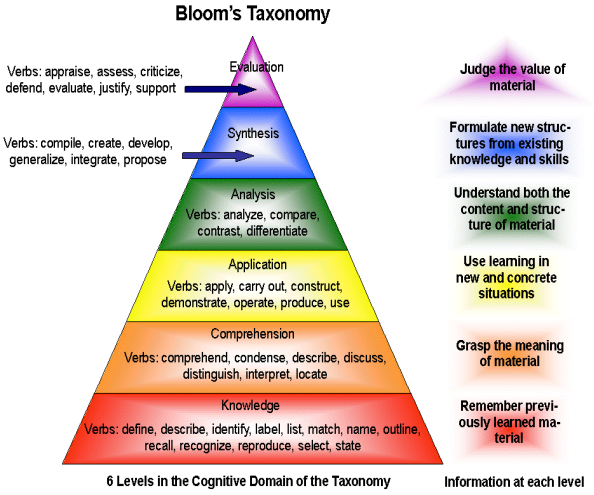

Figure 1: Cognitive Domains of Bloom's Taxonomy

Oral exams thus cover several cognitive domains and also the psychomotor skill of oral expression. In practice, oral exams are used not as substitute, but as a complement to written exams. Oral Examination enables interactive dialogue between candidate and assessor, allowing the examinee to demonstrate strengths between superficial and real knowledge via in-depth questioning and also the ability to modify the questions asked to the needs of each individual candidate.[15] This form of assessment is well suited for the evaluation of reflective and critical thinking competencies along with problem-solving abilities and analytic abilities. Oral exams also have the potential to measure the student's

achievement in course outcomes not restricted to knowledge, but related to individual's professionalism, ethics, interpersonal competence and qualities.[16]

Recommendations for increasing the effectiveness of oral exams

Orient the students

The candidates should be informed about the examination process in advance. The recommended preparatory techniques such as guidance from the supervisor, clearly defined guidelines, and mock oral exams can reduce students' stress levels.[17]

Train the examiners

Examiner performance can be enhanced by appropriate guidelines and instructions and training of new examiners. It may produce more uniform delivery of questions and evaluation of performance. Formal training programs can be adopted to increase both validity and reliability of oral examinations. Despite concerns about the subjectivity of oral exams, a longitudinal study of oral practice examination within medical programs revealed substantial internal consistency and reliability of orals, identifying a positive correlation to in-training examination scores and evaluation scores.[18, 19]

Use multiple assessors

Norman suggested that the oral examination must sample more broadly across cases and examiners to enhance reliability and scope of feedback.[20]

Assess on multiple occasions

Use a number of orals to enhance reliability and perceptions of fairness and accuracy

Questions should be straightforward and clear

Questions should be capable of being asked in a few sentences which are clear, unambiguous, uncomplicated, and without repetition.

Use simple grading system or rubrics

Criteria for answers can provide clear guidelines on what is not an acceptable answer to the examiner's questions. Checklists have been suggested as 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".[21] Examiners ratings of each student can be summated to give a score out of 60, which can then be converted to a percentage contribution reflecting his or her performance for this assessment component.[22]

Structure the oral on clinical scenarios

Structured oral examination (SOE) based on a clinical case with well-defined goals can often give great insight into a candidate's knowledge, interpretive ability, problem solving and attitudes thereby improving the inter-rater reliability results. Most authors agree that SOE has better validity and reliability, with less susceptibility to gender or cultural bias than unstructured examinations.[23]

Establish quality assurance standards

It is highly recommended to implement standards, benchmarks, performance indicators for effective oral examinations.[24]

CONCLUSION

Single assessment does not fulfill all aspects of assessment and there is a need for an evaluation system with multiple ways of assessment. Examiners differ in their personality, style and level of experience with variation of questioning and scoring from student to students. So, weakness of reliability on the oral examination still exists. Now-a-days worlds' leading medical schools used oral examination only for borderline and distinction students. Considering all these challenges, current assessment practices would be enhanced if the following principles summarized below were implemented:

- The content, format, and frequency of assessment, as well as the timing and format of feedback, should follow the specific goals of the medical education program.
- Proper orientation and faculty developmental activities are essential.

- The various domains of competence should be assessed in an integrated, consistent and longitudinal fashion with the use of multiple methods and provision of frequent constructive feedback.
- Educators should be mindful of the impact of assessment on learning, the potential unintentional effects of assessment, the limitations of each method (including cost), and prevailing traditions of the program or institution in which the assessment is occurring.
- Multiple methods of assessment implemented longitudinally can provide the data that are needed to assess trainees' learning need and to identify and remediate suboptimal performance by clinicians.
- The consistency and comparability of assessments methods should be maximized through a quality assurance system.
- Educators also face the challenge of developing tools for the assessment of qualities such as professionalism, teamwork, and expertise that have been difficult to define.

Conflicts of Interest There is no conflict of interest.

REFERENCES

1. Black P and Williams D. "assessment and classroom learning."Assessment in Education. 1998; 5(1): 7-74.
2. Epstein RM and Hundert EM. Defining and assessing professional competence. JAMA. 2002; 287: 226-235.
3. Leach DC. Competence is a habit. JAMA. 2002; 287: 243-244.
4. Klass D. Reevaluation of clinical competency. Amm J Phys Med Rehabil. 2000; 79: 481-486.
5. Epstein RM. Assessment in medical education. N Engl J Med. 2007; 356: 387-396.
6. Joughin G. Dimensions of oral assessment. Asses Eval High Educ. 1998; 23: 367-387
7. Kellough RD and Kellough NG. Secondary School Teaching: A Guide to Methods and Resources; Planning for Competence. 1999. Prentice Hill, Upper Saddle River, New Jersey.
8. Amin Z. Purposeful assessment. Medical Education. 2012; 46: 3-12. http://dx.doi.org/10. 1111/j. 1365-2923. 2011. 04170. x PMid: 22150188.
9. Salam A, Rabeya Y, and Mahmood CB. Assessment in Medical Education. JCMCTA; 15 (1&2): 15-18.
10. Van der Vleuten C. The assessment of professional competence: Developments, research and practical implications. Adv Health Sci Educ. 1996; 1: 41-67.
11. Newbel DA. Handbook for Medical Teachers. In: Norwell, 2nd ed. Boston, Massachusetts: MTP Press: 1987.
12. Memon MAG, Joughin and Memon B. "Oral assessment and postgraduate medical: establishing conditions for validity, reliability and fairness". Advances in Health Sciences Education. 2008.
13. Rousson V, gasser T and Seifer B. Assessing intrarater, interrater and test-retest reliability of continuous measurements. Stat Med. 2002; 21: 3431-3446.
14. Bloom BS. Taxonomy of Educational Objectives, Handbook I: The Cognitive Domain. New York: David McKay Co Inc; 1956.
15. Gibbs H, Habeshaw S and Habeshaw T. Interesting Ways to Teach: 53 Interesting Ways to assess your students. Bristol: Technical and Educational Services; 1988.
16. Harden RM. Developments in outcome based education. Med Teach. 2002; 24: 117-120.
17. Tinkler P and Jackson C. In the dark? Preparing for the PhD viva. Qual Assur Educ. 2002; 10: 86-97.
18. Birley H. The Society of Apothecaries Diploma Examination in Genitourinary Medicine: death of the viva voce? Sex Transm Infect. 2001; 77: 223-224.

19. Iqbal IZ, Naqvi S, Abeysundara L and Narula AA. The value of Oral Assessments: A Review. Bull R Coll Surg Engl. 2010; 92: 1-6.

20. Norman G. Examining the examination: Canadian versus US radiology certification exam. Can Assoc Radiol J. 2000; 51: 208-209.

21. Muzzin LJ and Hart L. Oral examinations. In: Neufeld RV, editor. Assessing Clinical Competence. New York: Springer Publishing Co; 1985.

22. Pearce G and Lee G. Viva voce (oral examination) as an assessment method insight from marketing students. J Mark Educ. 2009; 31: 120-130.

23. Simpson RG and Ballard KD. What is being assessed in the MRCGP oral examination? A qualitative study. Br J Gen Pract. 2005; 55: 430-436.

24. Morley L, Leonard D, and David M. Quality and equality in British PhD assessment. Qual Assur Educ. 2003; 11: 64-72.