Evaluation of structured oral examination format used in the assessment of undergraduate medical course (MBBS) of the University of Dhaka

Md Shah Alam[1] and Tahmina Begum[2]

[1]Department of Forensic Medicine, Ibrahim Medical College, Dhaka; [2]Department of Paediatrics & Neonotology, Bangladesh Institute of Research and Rehabilitation in Diabetes, Endocrine and Metabolic Disorder (BIRDEM), Dhaka

Abstract

Objectives of this cross sectional descriptive study was to evaluate critically the current status of structured oral examination (SOE) format as practiced in the professional examination of undergraduate medical course (MBBS) and views of the faculties regarding the concept of SOE as an assessment tool.

The study was conducted in 9 medical college examination centers of Dhaka University in July 2007. There were 36 examiners in 18 SOE board, 26 of them were interviewed with a semi-structured questionnaire and SOE boards were observed with a checklist. A total of 2455 questions used in SOE to assess 123 students, were recorded and analyzed using another checklist. These questions were used to assess learning hierarchy and content coverage using forensic medicine as a reference subject.

Analysis of the questions revealed majority (97%) were of recall type, only 3% were interpretation and problem solving types. The questions for 119 (97%) examinee did not address 10%-50% content area. About 38% examiners responded that they had no clear idea regarding learning objectives and none had idea regarding test blueprint. The examiners marked the domain of learning measured by SOE in favor of cognitive skill (61%), communication skill (38.5%), motor skill (11.5%), behavior and attitude (19%). No examiner prepared model answer of SOE questions by consensus with other examiner. Though more than 80% examiner agreed with the statement that pre-selection of accepted model answer is an important element for success of SOE. But no examiners of any SOE boards practiced it. Similarly, none of the examiners of SOE board kept records of individual question and the answer of the examinees. No boards maintained equal time for a candidate during SOE by using timer or stop watch. Examiners of 8 boards (44%) did not use recommended rating scale to score individual response of examinee rather scored in traditional consolidated way at the end of the candidate's examination. Majority (94%) boards scored the prompted answer and allowed another questions when a candidate failed to answer. During SOE conduction, 22% examiner were absent from the board for a prolonged period and 3% was engaged in marking the written scripts. About 56% of the examiners arrived late than schedule time. Behaviors of 14% examiner showed abusing to the candidates.

The study revealed that the objectives of introducing SOE as assessment tool in undergraduate medical curriculum was not achieved and it was not appropriately implemented. The various elements of SOE were not followed in most of the sessions of examinations. However, the reasons for not implementing vis a vis following the attributes of SOE were not explored. The study was done only in forensic medicine but similar situations may exist in other subjects. It is recommended that further study may be instituted to determine the causes of not achieving the objective of SOE in undergraduate medical evaluation system. The examiners should be motivated and trained up adequately to implement the elements of SOE successfully as valid, reliable and objective assessment tool.

Address for Correspondence:
Prof Md Shah Alam, Professor & Head, Department of Forensic Medicine, Ibrahim Medical College,122 Kazi Nazrul Islam Avenue, Shahbat, Dhaka 1000. Email: drpanna@yahoo.com

Introduction

Oral examination is traditionally an integral part of the evaluation of the undergraduate medical education. Oral communication dominates most fields of professional practice; therefore oral assessment is authentic in that it replicates the context of professional practice[1]. For this reason oral assessment is well established in Medicine, Law and Architecture.

Goal of assessment is to provide direction and motivation for future learning and protect the public by upholding high professional standards and screening out trainees and physicians who are incompetent. Learning abilities must be assessed in multiple modes and contexts. Educational contents are the stimulus for learning and also provide a context to demonstrate one's ability. Attributes of any instruments to assess different learning outcome should have four factors- validity, reliability, objectivity and practicality. The preferred learning style may be modified depending on the student's perception of task and motivation towards it.[2] Students' learning is influenced greatly by the assessment method used[3].

The major change in the way of assessment of medical students from 1950 to 1999 was the decision to replace essay question with MCQ format. Questions about credibility, reliability and validity of essay question in medical education lead to replacement by MCQ format in 1950s both in USA and UK.[4,5] The other examination format that dominated in the first part of last century was the oral examination. Essay and oral examinations are still popular in UK and other European countries though excluded for more than 20 years from assessment in North America on the ground of unreliability. Orals are unreliable due to lack of standardization of questions, insufficient judge and lack of sufficient time.[6,7] Orals can be highly threatening for candidates with resultant poor performance.[8]

The cognitive ability is assessed by the written examinations like essay question, modified essay question, short essay question (SAQ) and MCQ while skill by practical demonstration (OSPE/OSCE). The oral examination is still used in all subject centered medical curricula. Compared to essay questions, it is considered to probe deeply a student's ability to think, to express more or less clearly his knowledge of isolated facts or group of facts that he ought to remember. For the measurement of these reasoning and deductive process, problem solving skill, capacity to defend decision, evaluation of competing choice and ability to prioritize, still make the oral examination a popular tool in summative assessment. Oral examination has its unique characteristics as face to face interaction, flexibility to concentrate on one area and exploration of the student view points.

The use of oral examination has been criticized because of low reliability that relates in part to the active participation of the examiner which may introduce bias.[9,10] The candidate may receive a different assessment with regard to content areas addressed, the difficulty of the question asked, the level of prompting or help provided and the learning outcomes assessed. The reason for low reliability has an impact on validity because of the potential for variation in content areas addressed and in emphasis given on different areas. [11] A test blueprint may be used to ensure in obtaining the desired coverage of topics and level of objective to assess. The students from ethnic minorities and those trained from abroad experience hidden difficulties in language and thus discrimination as has been reported in MRCGP examination.[8] Female also get discriminated.

In view of those shortcomings, the oral examination should only be used to test the qualities that cannot be assessed by other method of evaluation. The qualities that are needed as medical professional include: Alertness, Confidence, Decisiveness and Ability to discuss logically.

There was evidence that the structured oral examinations were more reliable than the unstructured examinations.[12] A comparative study of traditional oral examination (TOE) and structured oral examination, found that the SOE is more effective, skillful technique and superior than TOE in assessing student's competency and cognitive ability.[13]. Structured oral examination based on clinical case with well defined goals can often give great insight into the candidate's knowledge, interpretative ability, problem solving ability and attitude.[14] One study concluded that the SOE can best evaluate the elements of problem solving.[15] A study in Bangladesh reported that about 81% of the teacher had opinion that there is no chance of subjective evaluation of any of the student by structured oral examination.[16]

In the unstructured oral examination, the examinees are liable to be asked whatever the examiners chooses and there is a risk that the examiner may concentrate

on his pet interests.[17] Assessment of medical students using the traditional oral system has been marred by being highly subjective, non-structured and biased; and therefore suggestion was for the replacement the traditional oral examination by the ViPSCE for testing knowledge, problem solving and management abilities.[18]

To overcome the limitations and to improve the oral examination, the SOE has been developed to make the assessment objective and structured. In SOE system, question of varying degrees of difficulty, rating scale, and correct answer are prepared prior to the examination which are discussed between the examiners and decisions are made jointly as to what is expected from the examinee. Equal time is allotted to each candidate by using timer or stop watch and a cordial atmosphere of SOE is ensured where the candidates feel secured and tension free environment. Assessment guidelines and model answer or list of criteria in SOE system help to neutralize the assessor effect to some extent.[19]

In the undergraduate medical curriculum of 2002 of Bangladesh, extensive modification of the assessment system was done. In this new curriculum the written examination format modified to SAQ and MCQ along with 10% mark added by formative assessment. Traditional practical and oral assessment were modified to OSPE/OSCE and SOE. The curriculum recommends that while constructing the questions for SOE, the proportion of recall, interpretative and problem solving questions should be 50-60%, 20% and 10-20% respectively. Questions should be constructed by the examiner and typed on a card for the candidate to pick up the cards randomly from a box. Two boards consisting four examiners should conduct the examination. Each candidate is allotted fifteen minutes to answer.

The SOE system was implemented in the pre- and para- clinical MBBS professional examinations in January-2005 and in January-07 respectively. In Bangladesh 4646 medical students are enrolled in each year by 15 public (2120) and 25 private (2526) medical colleges.[20] It is generally observed that being a new system many departments have yet not been able to implement the method with ease and up to the desired level of perfection. Considerable level of difficulty is expressed both by the examiners and the examinee. As the SOE system is universally accepted method now and one of the yardsticks to measure the quality

of medical education. The examiner must have an understanding of the assessment process as well as in designing and implementing this new program. The use of this system to make wider by making it easier and error free, it is necessary that the shortcoming in the present use of the SOE is understood correctly and remedial measures undertaken. Teachers hold strong personal opinion on testing, fostered by their own educational history and experience. Research outcome that demonstrate the opposite of such naïve intuition are often not immediately accepted. The assessment is dominated by tradition and intuition rather than research outcome[12]. Hence, this study was designed to find out the current practice of SOE undergraduate medical course (MBBS)in Bangladesh.

Materials and Methodology

This Cross sectional type of descriptive study was carried out over one year period from July 2007 to June 2008. The study was carried out in nine medical college centers under the University of Dhaka during the SOE of forensic medicine. The SOE in forensic medicine subject was chosen for evaluation as a reference and a representative discipline from other eleven subjects of MBBS course.

The sample size was 18 SOE boards (2 boards in each center), 26 examiners, 123 examinees and 2455 questions. The data collection instruments used in this study was 2 checklists and a questionnaire. Necessary modifications were made by a pre-testing as well as repeated testing to ensure that the response and finding were consistent.

Prior permission to observe was taken from appropriate authority. The conveners of examinations were informed about the intension for observing the session. The examiners were ensured that there would be no interference other than observation, documentation and collection of question asked to each candidate.

A checklist was prepared covering the key issues of SOE bearing in mind the objectives of the present study which contained 20 statements. Researcher himself with a lecturer of Forensic Medicine observed the SOE boards and collected all the data in real time using the 1st Checklist. There were two boards in the professional examination, each boards of examination consist of one internal and one external; one of the boards attended by the researcher himself and another by his colleague. Using checklist the observers indicate

whether or not a specific behavior or task were completed during the SOE. The semi-structured questionnaire was distributed to the examiners either before or after the examinations. The questionnaire containing 14 statements related to the elements of SOE practiced by the examiner to assess a candidate and opinion regarding certain aspect of SOE. These 14 statements of which 12 were close ended and 2 were open ended that focused on the objectives of the study. These two open ended questionnaire along with other 6 from closed ended questions, allowed the respondents the scope of giving free and extended comments and reason about the basis of their opinion. The close ended questions were of Yes/No type to ensure uniformity of measurements. During responding the questionnaires the researchers accompanied the respondents the whole duration and explained the educational issues that were not clear to them. The interview explored the examiners knowledge and practice as well as attitude, perceptions and preferences regarding the elements of SOE as an assessment tools in summative examination. The researcher analyzed the domain of the questions using the 2nd checklist. Inconsistency and uncertainty of data was removed in the field after collection of data before analysis.

Analysis was done using descriptive statistic.

Results

There were 36 examiners, of them 26 were included in this study. The demographic characteristics of examiners showed 24 (92%) were male and 2 (8%) were female and among them 11 (42%) were working in public medical colleges and 15 (58%) in private medical colleges. The respondents had mean teaching experience of 12.6 years and mean experiences as examiner of different universities of 9.3 years (Table-1).

Characteristics of a total of 2544 SOE questions that were asked to 123 examinees in relation to their learning hierarchy and content area in the forensic medicine showed majority (97%) of the questions were of recall type and very negligible numbers were interpretation (2%) and problem solving (1%) types (Table-2).

Highest percentage of questions were from Forensic Pathology (24%), followed by 22.5% from Forensic Toxicology, from Forensic Gynecology (12%) and Forensic Thanatology (11%). About 7% of questions were from Introduction and Legal Procedure and 6%

Table-1: *Distribution by characteristics of respondent examiner*

Designation	Sex		Experience as examiner Mean yr	Teaching experience Mean yr	Institution	
	Male	Female			Public	Private
Professor	12	-	15	17	05	07
Associate Professor	04	-	08	14	01	03
Assistant Professor	08	02	03	07	05	05
Total	24	02	9.3	12.6	11	15

from Medical Ethics. The detail distribution is given in Table 2.

Among 123 candidates only 4 (3%) were asked from the total content area of forensic medicine and 119 (97%) were lacking content ranging from 10% to 90% (Fig-1).

A significant number of examiners (10,38%) mentioned that they had no idea regarding learning objectives.

All boards (18) used the questions in strips that candidate picked up from the box and the examiners also asked some spot questions out of the strip to the examinees. No card was framed to include all 10 questions for a candidate with distributing the learning hierarchy and content. None of the boards used strip of written topic and the examiner could ask any questions from that topic area (Table-3).

None of the examiners of 18 boards practice to prepare model answer nor even recorded the questions that

Table-2: *Distributions of the SOE questions by their content area in Forensic Medicine (n=2544)*

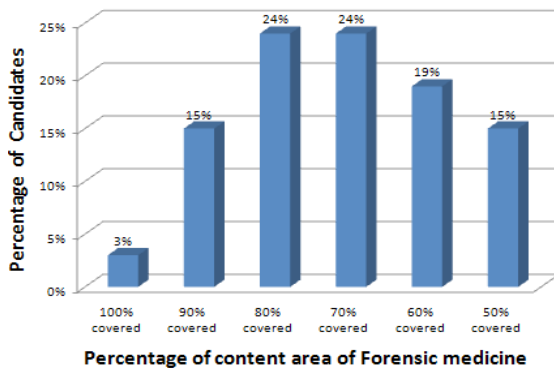| Topics	%	Recall	Interp- retative	Problem solving
Introduction to FM & Legal procedure	7			
Identification	9			
Medical Ethics & DPR	6			
Forensic Thanatology	10.7			
Forensic Pathology	24.2	2467	46	33
Forensic Gynecology	12.3	(97%)	(2%)	(1%)
Forensic obstetrics	5.6			
Forensic Psychiatry	2			
Forensic toxicology	22.5			
Forensic science, Trace evidence, domestic violence, Torture	0.4			

Fig-1. *Distributions of the candidates by the SOE questions they were asked from the total content areas of Forensic Medicine (n=123)*

Table-4: *Distribution of SOE by their procedure of conduction by examiner(n=18)*

Issues of the Examination Procedure	Yes n	No n
Earlier preparation of accepted model answer	0	18
Recording of question asked to each candidate	0	18
Recording of answer given by each candidate	0	18
Use of rating scale for scoring every	10	08
Scoring practice not using the rating scale concurrently with candidates answer but in a traditional consolidated way at end	08	10
Scoring of prompted correct answer	17	01
Allowed to shift if fail to answer in one area	17	01
Allowed equal time for each candidate by maintaining stop watch	0	18

were asked to the candidates and the answers of those questions. The examiners of 10 boards (56%) practiced scoring of every answer using rating scales but 8 (44%) of them scored in traditional consolidated way at the end when the candidates completed answering. Majority of examiner of the boards (94%) practiced scoring of prompted answer and equal number shifted to another question when candidates failed to answer. Equal time for a candidate was not maintained by stop watch in any of 18 boards (Table 4).

The atmosphere of SOE in 18 SOE boards revealed that 86% examiners were cordial to the examinees but 14% behaved non-cordially even in a abusing manner to the examinees (Table 5).

Other aspects of the behaviours of examiners during the SOE conduction revealed that 6-28% was involved in other activities during the examination procedure (Table 5). Only 28% of internals arrived one hour before schedule time. Both the internal and external

talked over cell phone and eating food during the SOE (Table 5).

Only 11 (42%) examiners prepared questions for the SOE, remaining 58% did not prepare questions for current session. Only 4 (15%) examiners mentioned that they use test matrix as a guideline for construction of question to consider hierarchy of learning. All 26 (100%) examiners did not practice prior selection of the model answer with consultation among them (Table-6).

The opinion of the examiners regarding the domain of learning outcome measured by the SOE, was a multiple response type question, hence 11 examiners marked more than one area. The opinions were 61% in favor of cognitive skill, 38.5% in favor of communication skill,

Table-3: *Pattern of framing sets of question used in SOE boards (n-18)*

Pattern of framing					No of set of questions	
All ten questions framed in one card	Individual question in strip	Question asked on spot*	Spot & from strip	Topic picked up & spot question asked from that topic	More than one set**	Only one set
0	18	0	18	0	0	18

* None prepared question earlier.
** One for examiner & other for candidate

Table-5: *Distribution of(atmosphere of SOE) behavior of examiner(n=36)*

Behavior of examiner	Internal Examiner(n=18)		External Examiner(n=18)	
	Yes n(%)	No n(%)	Yes n(%)	No n(%)
Cordial attitude	18(89)	02(11)	15(83)	03(17)
Abusing Behavior	02(11)	16(89)	03(17)	15(83)
Marking script	00(00)	18(100)	01(06)	17(94)
Telephoning	18(100)	00(00)	18(100)	00(00)
Eating	18(100)	00(00)	18(100)	00(00)
Prolong outside stay	04(22)	14(78)	06(33)	12(67)
Timely arrival	18(100)	00(00)	10(56)	08(44)
Arrival 1 hour before schedule time	10(56)	08(44)	00(00)	18(100)
Departure	18(100)	00(00)	18(100)	00(00)

Table-6: *Distribution of the respondent regarding their practice on SOEs (n=26)*

Aspect of SOE	Teacher's Practice	
	Yes	**No**
Prepared question for SOE	11(42%)	15(58%)
Used test blueprint as a guide for construction of question	04(15%)	22(85%)
Prepared accepted model answer for the questions of SOE by discussion among the four examiner	0	26 (100%)

11.5% in favor of motor skill and 19% in favor of behavior and attitude. Only 11.5% teachers had no idea regarding measurement of learning domain (Table 7).

More than 80% of examiners agreed (SA-38.5% & A-42.3%) that pre-selection of model answer was essential as prior structured questions for the success of SOE but 15 % disagreed. Only 4% examiners could not decide in selecting any one of the options Table-8.

About 31% of examiners agreed that test blueprint provide a ground rule for construction of the question of learning hierarchy' while 62% could not decide in selecting any one of the options (Table-8).

Advance construction of structured questions were preferred by the examiner as 1st highest priritorization. 2nd highest prioritization was in favour of advance preparation of model answer. 3rd highest prioritization was in favour of advance prepartion of model answer with consensus among the examiner. Creating nonthreatening environment during SOE was prioritize as 4th highest. 5th highest prioritization was in favour of providing 'equal time for each candidate'. 6th highest prioritization was 'providing equal time' and 'instruction to examiners about SOE'. Recording of questions and answers were also prioritize.

Highest net priority score (149) was in favor of the advance construction of SOE questions followed by advance preparation of model answer (127). The next

Table-7: *Teachers' opinion by the domain of learning outcome they want to measure by SOE (n-26, multiple responses)*

Cognitive skill	Communication skill	Motor skill	Attitude	Undecided
16(61%)	10(38.5%)	3(11.5%)	5(19%)	3(11.5%)

Table 8: *Distribution of teachers by their opinion about test blueprint and model answer for SOE (n=26)*

Item	Strongly Agree	Agree	Undeci-ded	Dis-agree	Strongly disagree
Preselection of model answer of the questions are essential for success of SOE	38.5%	42.3%	4%	7.7%	7.7%
Test blueprint is essential for construction of the questions of learning hierarchy & content coverage	8%	23%	62%	7.6%	

priority score was for non threatening environment (106), use of rating scale (95) and equal time for each candidate (93). Priority score for recording of question and answer was negative (-36) (Table 9).

Examiners opinion about advantages, disadvantages and suggestions for improvement of SOE were diverse in nature so those were analyzed qualitatively after grouping them into different categories. Some examiners opinions fell into two or more categories. The detail advantages and disadvantages of SOE is shown in Table 10 and 11. But the reasons for their

Table-9: *Distribution of elements of SOE by their net priority*

Essential elements of SOE	Net Score
Advance construction of structured question	149
Advance preparation of model answer in consultation with other examiner	127
Ensuring non-threatening environment during conduction of SOE	106
Scoring each answer by using rating scale	95
Careful selection & instruction of the examiner regarding SOE	58
Each candidate allow equal time(10 minutes) to answer by using stop watch	93
Recording of the questions & the answers of the candidate	-36

Note: *Net score of priority preference of essential elements of SOE was calculated by adding total score of essential element preferred by the examiner minus total score of non-essential elements. The scoring was allotted as: when the examiner preferred 1st score was 7, if preferred 2nd then score 6, if preferred 3rd then score 5, if preferred 4th then score 4, if preferred 5th then score 3, if preferred 6th then score 2 and if preferred 7th then score 1*

Table-10: *Distribution of the respondent by their identified advantage of SOE*

Characteristics	n(%)
Much better than traditional oral*	26(100)
Good system for rational and proper judgment*	26(100)
Valid, reliable tool of assessment*	24(92)
Equal time for each candidate can be ensured	26(100)
Unbiased & uniformity in scoring for all students	26(100)
Entire subject can be assessed ensuring learning hierarchy	22(85)
Disagreement among examiner can be minimized	26(100)
Time management is possible	26(100)
Nonthreatening environment can be ensured	26(100)

opinion was not mentioned by the respondents. All examiners identified disadvantage of SOE as misconception hinder the success, time not sufficient, laborious and time consuming, repeated use of same questions decreases the quality (100%) and success depends on luck (88%).

Regarding improvement of SOE, the respondents suggested in favour of regular training and workshop program for faculty, construction and upgrading of SOE question every year preparation of model answer, frame all the question in a card to one candidate, review of question and brief instruction for the examiner (Table

Table-11: *Distribution of the respondent by their identified disadvantages of SOE*

Characteristics	n(%)
Time is not sufficient to assess a candidate*	26(100)
Laborous & time consuming for examiner**	26(100)
Assess superficially; No provision for spot questioning	20(77)
Role of examiner during conduction of SOE is less important**	26(100)
Quality will detoriate,if same questions used repeatedly	26(100)
Success depends on luck**	23(88)
Most of the examiner are not engaged in construction of question & upgrading	20(77)
Misconceptions about SOE are hindering its success	26(100)

Time allotted for SOE is 15 minutes in all subjects.But in Forensic medicine 10 minutes for SOE & 5 minutes for traditional practikcal
*** Reason not mentioned.*

Table-12: *Distribution of respondent by their suggestion for improvement of SOE*

Characteristics	n(%)
Regular training & workshop for faculty development regarding educational technologies & SOE	26(100)
Preparation, standardization & updating of model answer centrally by a group of faculty	26(100)
Construction & updating of the questions yearly	26(100)
Framed all 10 questions in a card ensuring learning hierarchy& core content; Keep in a question bank centrally & distributed to each center.	22(85)
Review set of question(card) regarding learning hierarchy,content coverage & time needed to answer to correlate schedule time	26(100)
Review of the poor performer in SOE with good formative score	20(77)
Faculty must be instructed throughly before each SOE	
Misconception leads to misuse of SOE and will fail if all elements of SOE are not strictly maintained.	26(100)
Misuse of the system occurs if the examiner are not careful in implementing all the elements of SOE	22(85)

12). Review was sugested by 22 (77%) examiners of poor performer in SOE with good formative score.

Discussion

The characteristics of questions by their learning hierarchies and content coverage did not confirm the practice according to curriculum criteria. The majority questions were of recall type (97%) instead of 50-60% as specified in curriculum. Only negligible numbers (3%) were interpretation and problem solving type. Assessment in medical education must validate the objectives set by curriculum for three domains namely knowledge, skill and attitude. The knowledge includes all the cognitive process from mere recall through comprehension and understanding to an ability to solve problem; Assessment programs must match the competencies being learnt and teaching format being used.

The low taxonomic level (recall of factual knowledge rather than problem solving) of this study indicates, the students were adopting surface approach in learning. The study of learning style of medical students had

high scores on reproducing orientation were the evidence of surface approach in learning style.[21] The preferred learning style may be modified depending on the students perception of task and motivation towards it. Student learning is influenced greatly by the assessment method used.[3] Assessment strategies that focus predominantly on recall of knowledge will likely promote superficial learning but assessment strategies that demand critical thinking or creative problem solving will promote higher level of student performance or achievement. Higher education institution have been responding to a growing concern for the adequacy of professional and career preparation by specifying the outcome or abilities critical for future professional performance. Recent developments in assessment methodology have focused on performance assessment and good assessment can help students become more effective self-directed learners.[4]

Another important finding revealed by this study was wide variation regarding content area. Ninety seven percent of candidates appeared the SOE board was not assessed on content ranging from 10% to 50%. The potential for variation in content matter addressed and emphasis given to different content areas would definitely contribute to low reliability and also adversely affect validity.

Mainly internal examiners prepared the question for the current SOE session. This finding suggested reluctance of shouldering the responsibility by the examiner of SOE board. Wide variations in content and hierarchy discrimination could be avoided if the questions were constructed with the aid of prior prepared test specification (blueprint) and framing all 10 questions in a card distributing learning hierarchy and core content from all topics. Only 15% examiner responded that they used test blueprint in construction of question. Interestingly, even those examiners did not have proper knowledge about test blue print. All assessments should ensure that they are appropriate for the learning objective (Knowledge, skills and attitudes) being tested. The conceptual framework against which to plan assessment is essential and it is the test blueprint that provides a representative sample of instructionally relevant tasks. The test blueprint helps to achieve the validity of the content, response and consequence evidence.

About 58% of examiners categorized recording of questions and answers as essential elements of SOE.

Among them, 35% of respondent prioritized recording of questions and answer as 7th highest essential element of SOE, but no recording of question or response were practiced. The net priority score for this element was -36. None of the examiner realized the importance of recording; instead some examiner felt it unnecessary as well as time killing.

Scoring of students' response demands marking of every answer in a rating scale. To score in a traditional way at the end causes the subjectivity and bias. Examiners of 44% SOE boards did not use rating scale to score individual response rather they scored in traditional consolidated way at the end. It is quite impossible for an examiner to remember all the responses after a prolong time, without subjective bias. Therefore, scoring of all the answer traditionally at the end definitely invite bias in scoring.[22] The factors that influence rating, are the errors of leniency and central tendency and hallo effect which should be avoided in rating scale construction and use.[23]. But in this study no examiners prepared any rating scale and even significant portion of examiner scored traditionally at the end.

Examiners of any sessions of SOE did not maintain equal time for the individual candidate to answer by using stop watch, though all of them prioritized maintaining equal time as an essential element of SOE. The variation of time allocated to the candidate contributes subjective bias in scoring. This would contribute for low reliability and ultimately to validity of the assessment system.[9,10,24]

The respondent prioritized the advance preparation of model answer in consultation with other examiner's as 2nd highest essential element of SOE and majority examiner (80%) agreed that the pre-selection of model answer was important for success of SOE. The net priority score for preparation of accepted model answer was 127. But none of the examiner prepared any model answer of the questions. Specified answers and a specific marking scheme in an SOE for surgical resident in Canada produced an overall reliability of 0.75.[25] Criteria for answer can provide clear guidelines on what is and is not an acceptable answer to the examiner's question.

A widely recognized feature of oral examination is that it particularly focuses on the capacity to think quickly under pressure; therefore, measures to reduce stress should receive particular attention. A number

of non-verbal effects in an oral testing environment, such as head gestures and facial expressions are likely to vary greatly among staff.[31] It is mandatory to establish a cordial non-threatening environment where the candidates feel secured. Studies show that the oral examination can be highly threatening for candidates with resultant poor performance.[4] In the present study we found that 14% of the examiner's behavior was non-cordial or even abusing to the examinees. This negative attitude and abusing behavior was not acceptable and ethical but also contributed bias of subjectivity leading to low reliability and validity.[24]

The activities like talking over mobile, eating foods, prolonged outside staying and marking scripts during conduction of SOE, is unsuitable for establishment of cordial environment and unbiased scoring. Majority of external examiners were late than university schedule time. Examiners should arrived one hour before the university schedule time of starting SOE for selection of question, preparation of accepted answer on consensus and rating scale. However, it was interesting to note that all the examiners prioritized the establishment of a non-threatening environment as an essential element of SOE.

The present study attempted to determine the views of examiners regarding the concept (knowledge, practice and attitude) of SOE. The net priority score of essential elements of SOE, which were preferred by the examiners, were ranging from 58 to 149. The preferred essential elements were structured question (149), accepted model answer (127), non threatening environment (106), scoring individual question (95), equal time for each candidate (93) and instruction of examiner (58). Practically most of those elements were absent in practice. Though the demographic profiles of examiners indicate that they had adequate experience as university examiners (mean 9.3 years). Examiners did not prioritize recording of answers needed for future evaluation and bias free scoring as essential elements, even mentioned it unnecessary and time killing. However, further study is needed to explore the reasons of these discrepancies to take remedial measure.

Education is a process the chief goal of which is to bring change in human behavior. This behavior explicitly defined in the form of educational objectives, which are the guiding principles to plan educational activities and assessment. A significant percentage of examiners (38%) had no clear idea regarding learning objectives. The lack of knowledge regarding learning objectives indicate basic defect to overcome all the barriers of effective medical education system. Definition of educational objectives is an essential step before choosing teaching method and a system of evaluation. In the present study, about 12% respondent could not decide learning domain measured by SOE, 12% marked in favor of motor skill, 19% responded in favor of attitude and 10% for behavioral domain of learning. The findings were suggesting that significant number of examiners did not perceived the elements measured by SOE. The findings indicate that there is an urgent need of training for faculty development.

Advantages identified by the examiners if implemented would establish the objectives of SOE To avail those advantage, all the elements of SOE should be well understood by the faculty members involved in assessments. Some of disadvantage such as insufficient time for SOE, lack of provision for spot questing, element of luck, etc as identified by the examiner were due to misconception of SOE.

Conclusion

The elements of SOE were not properly followed during assessment of students' in forensic medicine. Without using test blue print in construction and framing of questions it is quite impossible to assess the candidate's learning hierarchy and coverage of essential content. Advance preparation of accepted model answer though essential in scoring without bias was not practiced by the examiners of any board. The medical colleges were selected purposively and therefore, all medical colleges could not be included. The examiners of different subject could not be interviewed and all SOE boards could not be observed However, the reasons for not implementing vis a vis following the attributes of SOE were not explored. The study was done only in Forensic medicine but similar situations may exist in other subjects also. The study revealed that SOE introduced as assessment tool in undergraduate medical curriculum was not properly implemented and its desired objectives not fully achieved.

Recommendation

The policy makers must take urgent necessary action to arrange regular and intensive training program for

faculty development. Further study may be undertaken to determine the reason(s) of not being appropriately implementing SOE and its suitability in our medical curriculum. Also, examiners should rethink for redesigning the SOE as an assessment tool.

References

1. Joughin G. Dimensions of oral assessment and student approaches to learning. In: Brown S, Glasner A eds. Assessment matters. Buckingham: The Society for Research into Higher Education & Open University Press 1999; 146-156.

2. Laurillard D. The process of student learning. *Higher Education*1979; **8**: 395-409.

3. Wood DF. ABC of learning & teaching in med problem based learning. *British Med J*.2003; **326**: 328-330

4. Pokorny AD & Frazier SH. An evaluation of oral examination. *J of Medical Education* 1966; **41**: 28-40

5. Bull GM. Examinations. *J Medical Education*. 1959; **34**: 1154-1158.

6. Swanson DB, Norcinim JJ, Groso LJ. Assessment of clinical competence: written & computer based simulation. *Assess Eval Higher Education*1987; **12**: 220-246.

7. Swanson DB. A measurement framework for performance based test. In: Hart IR, Harden RM eds. Further developments in assessing clinical competence. Montreal, CanHeal 1987.

8. Harper AC, Roy WB, Norman GR, Rand CA, Feightner JW. Difficulties in clinical skills evaluation. *Medical Education*. 1983; **17**(1): 24-27.

9. Colton T & Paterson OL. An assay of medical student's ability by oral examination. *J of Medical Education* 1967; **42**: 1005-1014.

10. Foster JT, Abrahamson S, Lass S, Girard R & Garris R. Analysis of oral exam used in specialty board certificate. *Medical Education* 1969; 44: 951-954.

11. Newble DI, Hoare J & Efmsli RG. The validity & reliability of a new exam of the clinical competency of medical students. *Medical Education* 1981; **15**: 46-52.

12. TuttonPJM & Glasgow EF. Reliability and predictive capacity of examination in Anatomy & improvement in the reliability of viva voce (oral) exam by use of Structured rating system. *Clinical Anatomy* 2005; **2**: 29-34.

13. Ferdousi S, Latif SA, Ahmed MM, Nessa A. Summative assessment of undergraduate medical students' performance in physiology by Structured Oral examination. *Mymensingh Med J* 2007; **16**(1): 64-69.

14. Fabb WE & Marchal JR. Assessment in clinical competence in general family practice (Lancester MTP press). 1983.

15. Kearny RA, Puchalski SA, Yang HYH and Skakun EN. The inter-rater and intra-rater of reliability of a new Canadian oral examination format in Anaesthesia is fair to good, Canadian. *J Anaesthesia* 2002; **49**: 232-236.

16. Khan TF. Implementing the new assessment system in undergraduate medical education in Bangladesh-teacher's view, (MMEd-Thesis)DU 2008; 74.

17. Oyebode F,George F, Math V & Haque S. Inter examiner reliability of the clinical parts of MRCPPsyc Part-II exam. *Psychiatry Bulletin* 2007; **31**: 342-344.

18. Shallaly GHE, Ali EA. Use of Video-projected structured clinical examination (ViPSCE) instead of the Traditional oral (Viva) examination in the assessment of final year medical students. *Education for Health*. 2004; **17**(1): 17–26.

19. van der Vleuten CPM, Scherpbier AJJA, Dolmans DHJ, Schuwirth LWT, Verwijnen GM, Wolfhagen HAP. Clerkship assessment assessed. *Medical Teacher* 2000; **22**(6): 592-600.

20. Amin Z, Merrylees N, Hanif A & Talukder HK. Medical education in Bangladesh. *Medical Teacher* 2008; **30**: 243-248.

21. Newble DI, Gordon MI. The learning style of medical students. *Medical Education*1985; **19**: 3-8.

22. Ghosh A, Mandel A, Das N, Tripathi SK, Biswash A, Bera T. Students performance in written & viva voce component of final summative pharmacology exam in MBBS curriculum: A critical insight. *Indian J Pharmacol* 2012; **44**: 274-285.

23. Guilbert JJ. Comparison of advantage and disadvantage of different types of test, Educational handbook for health personnel, Geneva, WHO 1977; 416.

24. Kelly PR, Matthews JH & Schumacher CF. Analysis of the oral examinations of the American Board of Anesthesiology. *J of Med Education* 1971; **46**: 982–988.