The Professional Medical Journal
www.theprofesional.com

DOI: 10.17957/TPMJ/17.3989

ORIGINAL PROF-3989

MULTIPLE CHOICE QUESTIONS;
DEVELOPED BY THE FACULTY OF A PUBLIC SECTOR MEDICAL COLLEGE

Muhammad Zafar Iqbal[1], Shumaila Irum[2], Muhammad Sohaib Yousaf[3]

1. MBBS, FCPS, MME
   Associate Professor
   Pediatric Surgery
   King Edward Medical College/
   Mayo Hospital Lahore.
2. MBBS, MME
   Senior Demonstrator,
   Physiology Department
   Sheikh Zayed Hospital
   Rahim Yar Khan.
3. MBBS (KEMU)
   House Officer
   Pediatric Surgery Department
   King Edward Medical University/
   Mayo Hospital Lahore.

Correspondence Address:
Dr. Muhammad Zafar Iqbal
MBBS, FCPS, MME
Associate Professor
Pediatric Surgery
King Edward Medical College/
Mayo Hospital Lahore.
drzafar300@yahoo.com

Article received on:
18/04/2017
Accepted for publication:
30/06/2017
Received after proof reading:
08/09/2017

ABSTRACT... Objectives: The main objective of this study was to judge the quality of MCQs in terms of their cognition level and item writing flaws, developed by the faculty of a public sector medical college. **Setting:** This study was conducted in Sheikh Zayed Medical College, Rahim Yar Khan. **Duration with Dates:** Data was collected between June 2014 to March 2015 and this study was completed in July 2016. **Sample Size:** A sample of 500 MCQs collected from 25 faculty members were included in the study. **Study Design:** Quantitative method. **Study Type:** Cross sectional descriptive analysis. **Material and Methods:** This quantitative study was conducted in Sheikh Zayed Medical College Rahim Yar Khan over six months period after the approval of the study proposal. Every faculty member is supposed to write 25 MCQs in order to become supervisor. I collected 500 multiple choice questions from 25 faculty members ready for submission to CPSP. The quality of all MCQs was checked in terms of item writing flaws and cognition level by panel of experts. **Results:** Absolute terms were observed in 10(2%), vague terms in 15(3%), implausible distracters in 75(15%), extra detail in correct option 15(3%), unfocused stem 63(12.6%), grammatical clues 39(7.8%), logical clues 18(3.6%), word repeats 19(3.8%), >then one correct answer 21(4.2%), unnecessary information in stem 37(7.4%), lost sequence in data 15(3%), all of above16(3.2%), none of above 12(2.4%) and negative stem 23(4.6%). Cognition level I (recall) was observed in 363(72.6%), level II (interpretation) in 115(23%) and level III (problem solving) in 22(4.4%) items. Total 378(75.6%) flaws were identified and four commonest flaws were implausible distracter 75(15%), unfocused stem 63(12.6%), grammatical clues 39(7.8%) and unnecessary information in stem 37(7.4%). **Conclusion:** It is concluded that assessment of medical students is very demanding and need of the time. A well-constructed, peer-reviewed single best type MCQ is best one to complete this task because of cost effectiveness, better reliability and computerized marking. It is very important to start faculty development program in order to decrease the number of item writing flaws and improve cognition level towards problem solving and application of knowledge.

Key words: Items, items writing flaws, cognition level, faculty development program.

Article Citation: Iqbal MZ, Irum S, Yousaf MS. Multiple choice questions; Developed by the faculty of a public sector medical college. Professional Med J 2017;24(9):1409-1414. **DOI:** 10.17957/TPMJ/17.3989

INTRODUCTION

Assessment is one of the very important components of teaching and learning. In other words, assessment drives learning and force students for better learning strategies.[1] So it is very important to find out the assessment technique which will help us to find out the approach of learning for students. It has been reported by many studies that methods of assessment influence the students approach towards learning.[2,3,4]

There are different tools available to be used for assessment in the field of medical education and most of them are based on three domains of educational psychology i.e. cognition, psychomotor and attitude. Among these tools, Multiple Choice Question (MCQ) is widely used in graduate and post graduate examinations because of its reliability and consistency.[5] Fixed choice, selected response or multiple choice items are other names used for multiple choice questions because they are not always questions. Sometime a scenario or case history is presented to students and they are asked to choose one correct option from various options that are offered to them. In literature there are many types of

multiple choice questions but the most frequently used are "A" and "R" type. In A-type MCQs, 4 or 5 different options are offered to students and they have to pick one best possible answer. It has also been observed that tests containing multiple choice questions often need less time to manage for a given amount of material than would tests consisting of written answers.

The anatomy of MCQs always consists of a scenario or a problem which is called stem. This stem is followed by a list of proposed solutions known as options or alternatives. The stem part of MCQs is relatively long and consists of as much information as possible but it should be remembered that there should be no unnecessary or unrelated information. Out of 4 or 5 alternatives, only one is correct and labeled as answer or best alternative. The other remaining options or alternatives which are incorrect or inferior alternatives are labeled as distractors.[6]

The presence of item writing flaws (IWFs) is most common hindrance in developing good quality items. These IWFs are different types of violations making the item either easier or sometimes more difficult and badly affect students' performance.[7]

For good quality items it is very important to avoid all structural flaws, this will uplift the validity of multiple choice questions as well as related examination.[8] Tarrant and Ware exclusively studied 10 nursing examinations and found that after removing flawed items 94.3% examinees were declared as passed which were otherwise 90.6%.[9]

Large numbers of items are developed by the faculty members of our newly established Medical College. These MCQs are used for assessment in MBBS examinations and are also the mandatory requirement by college of physicians and surgeons of Pakistan (CPSP) to become supervisor. The quality of these MCQs written by the faculty is not up to the mark. So I planned a study to find out the quality of items based on following two parameters:-

- Number of item writing flaws (IWFs)
- Cognition level of MCQ or items.

PATIENTS AND METHODS

Aims and Objectives

The main objective of this study was to judge the quality of MCQs in terms of their cognition level and item writing flaws, developed by the faculty of a public sector medical college.

Setting

This study was carried out in pediatric surgery department Sheikh Zayed Medical College/ Hospital Rahim Yar Khan, which is a tertiary care hospital.

Duration

Data was collected between June 2014 to March 2015 and this study was completed in July 2016.

Sample Size

A sample of 500 MCQs collected from 25 faculty members were included in the study

Sampling Technique

Non probability, convenience sampling

Study Design

Quantitative study

Study Type

Cross sectional descriptive analysis

Target Population

The target population for this study includes faculty members of all medical colleges of Pakistan who intend to become supervisors. These colleges are mainly medical and dental colleges and postgraduate training institutes.

Accessible Population

The accessible population for this study includes faculty members of Sheikh Zayed Medical College, who were requested to provide multiple choice questions for quality assessment.

Inclusion Criteria

Multiple choice questions prepared by the faculty members of Sheikh Zayed Medical College in order to become supervisor were included in the study.

Exclusion Criteria

MCQs prepared by teachers other than faculty members were excluded from study.

Data Collection Instrument

The whole data was collected based on two Performas. Performa 1 for fourteen possible item writing flaws and Performa 2 for item cognition complexity. These flaws are of two type test wiseness and irrelevant difficulty as under:-

Performa 1 (Item writing flaws)

	Test wiseness		Irrelevant difficulty
1	Grammatical cues	9	Vague terms
2	Logical cues	10	Lost sequence in numeric data
3	Word repeats	11	Unnecessary information in stem
4	Extra detail in correct option	12	Unfocused stem
5	Absolute terms	13	>1 or no correct answer
6	Implausible distracters	14	Negative stem
7	All of above		
8	None of above		

Performa 2 (Item cognition complexity)

MCQ Number	Level I Recall	Level II Application	Level III Analysis
MCQ 1			
MCQ 2			
MCQ 3			
MCQ 4			
MCQ 5			

500			

Data Collection Procedure

Data was collected from a total of 500 multiple choice questions (one best type). All MCQs were prepared by the faculty members from both basic and clinical departments. Apart from mandatory requirement to become supervisor, these items were developed for send-up examinations of all four proofs of MBBS classes. Best effort was done to get final copy of the questions. It was made possible by personal meeting with all faculty members and they were assured that these MCQs will be used for research purpose only. Each MCQ was keenly assessed by analysts (panel of experts) and number of possible flaws along with cognition level were identified and marked in a specially designed performs as mentioned above.

RESULTS

500 items from 25 faculty members belonging to basic and clinical departments were collected. These multiple choice questions (one best type) were banked from 22 different specialties in different quantity as shown in the table below.

Sr. no.	Specialty	Number of items
1	Pediatric surgery	40
2	General surgery	45
3	Gynecology	35
4	Biochemistry	15
5	Physiology	25
6	Pharmacology	10
7	Urology	20
8	Pathology	25
9	Orthopedic	20
10	Neurosurgery	18
11	Pediatric medicine	32
12	Medicine	30
13	Cardiology	25
14	Anesthesiology	22
15	Biochemistry	18
16	Nephrology	20
17	Radiology	14
18	Psychiatry	20
19	Plastic surgery	15
20	Eye	11
21	TBCD	25
22	Physiotherapy	15
	Total	500

Table-I. Specialty wise breakup of items

Following possible item writing flaws were found after thorough analysis of MCQs by panel of experts.

All MCQ were assessed for cognition level and cognition complexity was found as under.

Sr. No.	Possible flaws	Total questions	Flawed questions	%age	Remarks
	Absolute terms	500	10	2%	
	Vague terms	500	15	3	
	Implausible distracters	500	150	30	Maximum number of distracters
	Extra detail in correct option	500	15	3	
	Unfocused stem	500	135	27	Third commonest cause of flawed items
	Grammatical clues	500	65	13	
	Logical clues	500	35	7	
	Word repeats	500	45	9	
	>one correct answer	500	40	8	
	Un-necessary information in stem	500	145	29	Second commonest cause of flawed items
	Lost sequence in data	500	15	3	
	All of above	500	30	6	
	None of above	500	35	7	
	Negative stem	500	55	11	

Table-II. Item writing flaws

	Number of items	Percen-tage	Remarks
Level I Recall	363	72.6	Easy to prepare, at recall level only
Level II Interpretation	115	23	
Level III Analysis & problem solving	22	4.4	Prepared at synthesis and analysis level.

Table-III. Item cognition complexity

DISCUSSION

In Pakistan, appropriate selection of an assessment tool for measuring students' performance in medical colleges remains an intimidating task. Out of many assessment tools, "A" type MCQ is best one for identifying the strengths and weaknesses in student's performance. Apart from that single best type MCQs also provide guidelines to teachers on the basis of student's assessment report and help us to develop new educational protocols and strategies.[10]

The main target of any assessment technique is to test students' knowledge on the subject rather than their test-taking skills. For this purpose we need to get rid of as many flaws in our questions as possible. So the provision of a level playing field for both test wise and not-so-test wise students is mandatory requirement in any assessment method. The possibility of correct answering a question of a student should relate to his knowledge on the topic and should not relate to his test-taking strategies. Although lots of data is available on developing good quality questions but literature regarding analysis of MCQs is scare.

Our study results showed that the frequency of item writing flaws encountered was 75.6%. Out of 500 multiple choice questions 378(75.6%) were flawed. While it was 67% in a study conducted by Humaira Fayyaz Khan in 2013.[11] Apart from that when I compared results with other studies, the flaw rate was 75%[12] and 46.2%.[9] Similarly in another study conducted in Pakistan 69(46%) item writing flaws were observed in 150 items.[13] While in another study conducted by Ware J, item writing flaws were 31 out of 389(7%).[14] This is true that percentage of flawed MCQs is relatively high in our study as compared to other national and international studies, probably this is due to the new recruitment of whole faculty members and lack of proper training regarding items training.

All multiple choice questions contain one correct answer and three or four alternate answers. All alternate answers are referred as distracters. The implausible distracters were the commonest flaws observed in my study 75(15%). While in another study implausible distracters observed were 30.43%.[13]

Studies	Four common flaws			
Our study	Implausible distracters 15%	Unfocused stem 12.6%	Grammatical clues 7.8%	Un-necessary information in stem 7.4%
Mukhtair et al[13]	Implausible distracter 30.43%	Unfocused stem 27.54%	Un-necessary information in stem 24.64%	Negative stem 8.7%
Collin J [15]	Negative stem 17%(n=33)	Implausible distracter 14% (n=27)	extra detail in correct option 13%(n=26)	unfocused stem 10% (n=20)

Table-IV. Four commonest flaws

	Our Study		Mukhtair Baig[19]		Khan and Aljarallah[17]	
Cognition levels	Number	%age	Number	%age	Number	%age
Level I	363	72.6	114	76%	14	28%
Level II	115	23	36	24%	3	6%
Level III	22	4.4	0	0	30	60%

Table-V. Cognition level of different studies

Four commonest flaws observed in this study were implausible distracter (15%); unfocused stem (12.6%), grammatical clues (7.8%) and un-necessary information in stem (7.8%). When I compared my results with other studies, four commonest flaws observed were as under.

In my study negative stem was observed in 55 (11%) items while it was 8.7%[13] and 17%[32] in other studies. Similarly unfocused stem in my study was 27% while it was 10% in a study conducted by Collin J.[14] Un-necessary information in stem was 24.64%[13] while in my study it was 29%. So many types of structural flaws were observed in my study. The frequency of these flaws is different in different studies.

As for as cognition level is concerned, all 500 MCQs were checked by an experienced analysts. Level I was observed in 363(72.6%), level II in 115(23%) and level III in 22(4.4%) items. In another study conducted by Mohair et all[13] cognition level I was observed in 114 MCQs (76%), level II in 36(24%) and no MCQ was at level III. So in both studies problem solving domain of knowledge was not efficiently evaluated. It was due to the fact that MCQs at recall level are easier to construct and need less time and knowledge as compared to problem solving MCQs which needs more expertise and training.[15,16] In another study conducted by Khan and Aljarallah[17] the percentage of cognition level of MCQs was level III (60%), level II (6%) and level I (28%). In a nursing examination, Tarrant and Ware (2008)[18] found that item-writing flaws are significantly common in MCQs written at lower cognitive level.

MCQs for higher order cognition levels are the need of the time and faculty should be trained to construct. In 2006, Tarrant et al. stated that cognition level of items does not change by removing item writing flaws rather writing questions at higher cognitive levels inherently remove numerous IWFs.[20] The corollary is that in my study, multiple factors are responsible for low cognition level of questions such as recently established medical college, newly recruited teaching faculty and inadequate faculty training programs. Apart from that, diverse background of faculty members and casual attitude of teachers badly effected in developing quality items.

CONCLUSION & RECOMMENDATIONS

It is concluded that assessment of medical students is very demanding and need of the time. A well-constructed, peer-reviewed single best type MCQs is best one to complete this task. It is recommended that a well-structured faculty development program should be initiated for preparing high quality multiple choice questions. This outcome indicates that faculty development programs should be started on priority basis in a well-organized manner.

REFERENCES

1. Drew S. **Student perceptions of what helps them learn and develop in higher education.** Teaching Higher Educ. 2001; 6(3):309–331.

2. Trig well K, Prosser M. **Improving the quality of student learning: the influence of learning context and student approaches to learning on learning outcomes.** Higher Educ. 1991; 22:251–266.

3. Biggs J. **Teaching for Quality Learning at University Philadelphia: The Society for Research into Higher Education & Open University Press; 2003.**

4. Reid WA, Duvall E, Evans P. **Relationship between assessment results and approaches to learning and studying in Year Two medical students.** Med Educ. 2007; 41:754–762.

5. **Multiple choice questions: Their value as an assessment tool.** Moss E, Curr Opin Anaesthesiol. 2001 Dec; 14(6):661-6.

6. Kehoe, Jerard. **Writing multiple-choice test items Practical Assessment, Research & Evaluation, 4(9)1995.**

7. Downing S. M. **The effects of violating standard item writing principles on tests and students: the consequences of using flawed test items on achievement examinations in medical education.** Adv. Health Sci. Educ. 10, 133–143 (2005).

8. Downing SM. **Validity: on the meaningful interpretation of assessment data.** Med Edu. 2003; 37:830–7.

9. Tarrant M, Knierim A, Hayes SK, Ware J. **The frequency of item writing flaws in multiple choice questions used in high stake nursing assessments.** Nurse Educ. Today.2006, Vol.26, pp. 662-671.

10. Tan LT, Mc Aleer JJ; Final FRCR Examination Board. **The introduction of single best answer questions as a test of knowledge in the final examination for fellowship of the Royal College of Radiologists in Clinical Oncology.** Clin Oncol (R Coll Radiol) 2008; 20: 571-6.

11. Khan HF, Danish KF, Awan AS, Anwar M. **Identification of technical item flaws leads to improvement of the quality of single best Multiple Choice Questions.** Pak J Med Sci. 2013 May-Jun; 29(3): 715–718.

12. Hansen, JD. **Quality multiple-choice test questions: item-writing guidelines and an analysis of auditing test banks.** J Educ Bus.1997, Vol. 73, pp. 94–97.

13. Kehoe, Jerard. **Writing multiple-choice test items Practical Assessment, Research & Evaluation, 4(9)1995.**

14. Ware J, Vik T. **Quality assurance of item writing: During the introduction of multiple choice questions in medicine for high stakes examinations.** Medical teacher 2009, Vol. 31, No. 3, Pages 238-243.

15. Case S, Swanson D. **Constructing written test questions for the basic and clinical sciences.** Philadelphia: National Board of Medical Examiners; 2003.

16. Tarrant M, Ware JA. **Framework for improving the quality of multiple-choice Assessments**. Nurse Educator. 2012; 37(3):98–104.

17. Khan MZ, Aljarallah BM. **MEQs and MCQ as a tool for assessing the cognitive skills of undergraduate medical students.** Int. J Heal Sci. 2011; 5(1):45–50.

18. Tarrant M, Ware J. **Impact of item-writing flaws in multiple choice questions on student achievement in high-stakes nursing assessments.** Med Educ. 2008; 42:198–206.

19. Baig M, Ali SK, Ali S, Huda N. **Evaluation of Multiple Choice and Short Essay Question items in Basic Medical Sciences.** Pak J Med Sci. 2014 Jan-Feb; 30(1): 3–6.

20. Tarrant M, Knierim A, Hayes SK. **The frequency of item writing flaws in multiple-choice questions used in high stakes nursing assessments.** Nurse Educ Pract. 2006; 6:354–363.

AUTHORSHIP AND CONTRIBUTION DECLARATION

Sr. #	Author-s Full Name	Contribution to the paper	Author=s Signature
1	Muhammad Zafar Iqbal	Full	
2	Shumaila Irum	Full	
3	Muhammad Sohaib Yousaf	Full	