ARE POSTGRADUATE MEDICAL ENTRANCE EXAMINATIONS REALLY DIFFICULT? - ITEM ANALYSIS OF MCQs FOR DOCTOR OF MEDICINE IN INDIA

Poonam Kumari[1], Richa Niranjan[2]

[1]Assistant Professor, Department of Physiology, Government Medical College, Haldwani, Uttarakhand, India.
[2]Associate Professor, Department of Anatomy, Government Medical College, Haldwani, Uttarakhand, India.

ABSTRACT

BACKGROUND

Multiple Choice Questions (MCQs) are widely used for evaluations of knowledge for many professional courses including Medicine. Construction of appropriate test items is a challenge in preparing quality multiple choice questions. Item analysis provides valuable feedback data on validity of multiple-choice questions.

Aim- The present study was conducted to evaluate the Difficulty Index (DIF-I), Discrimination Index (DI) and Distractor Efficiency (DE%) of the multiple-choice questions (items) of previous postgraduate medical entrance examinations.

MATERIALS AND METHODS

A list consisting of 50 MCQs was taken from standard question bank (memory based) of Physiology, usually followed by postgraduate aspirants and administered to undergraduate students (n= 97). The result of student's performance in these MCQ based unit test was entered in MS Excel 2007. Each item was analysed for three parameters: Difficulty Index, Discrimination Index and Distractor Efficiency.

Study Design- Observational cross-sectional study.

RESULTS

In the present study, after analysis of data 28% of the MCQs were difficult, 20% were good, 26% were excellent and 26% were easy with Difficulty Index. 58% were excellent and 10% were good with Discrimination Index. 46% had at least one non-functional distractor.

CONCLUSION

Item analysis of MCQs of postgraduate entrance examinations demonstrated that Difficulty Index was neither too high nor too low. Discrimination Index was high, and Distractor Efficiency was good. Hence, we proposed that for formation of appropriate test items, question bank in the form of quality MCQs with negative marking should be prepared and their post validation, data to be collected.

KEY WORDS

MCQs, Item Analysis, Postgraduate Medical Entrance Examination, Difficulty Index, Discrimination Index, Distractor Efficiency.

HOW TO CITE THIS ARTICLE: Kumari P, Niranjan R. Are postgraduate medical entrance examinations really difficult? - item analysis of MCQs for doctor of medicine in India. J. Evolution Med. Dent. Sci. 2018;7(46):4956-4959, DOI: 10.14260/jemds/2018/1103

BACKGROUND

Multiple Choice Questions (MCQs) are widely used for evaluations of knowledge for many professional courses including Medicine.[1] Construction of appropriate test items is a challenge in preparing quality multiple choice questions. They are used to make decisions regarding the certification of candidate.[2] They identify strengths and weakness in students.[3]

This study focused on MCQs type of assessment, since in present scenario after obtaining MBBS degree (20 subjects in span of four and a half years plus one year of compulsory rotatory internship), all the aspiring graduates have to qualify competitive entrance examination based on MCQs for their desired speciality in medicine. The student's fate is decided on ranking in above said exam.

'Financial or Other Competing Interest': None.
Submission 16-10-2018, Peer Review 28-10-2018,
Acceptance 30-10-2018, Published 12-11-2018.
Corresponding Author:
Dr. Poonam Kumari,
Department of Physiology,
Government Medical College,
Haldwani, Uttarakhand, India.
E-mail: drpoonam99p@gmail.com
DOI: 10.14260/jemds/2018/1103

Several studies have been done for analysis of MCQs. Authors tried to find how "ideal questions" are framed, the relationship of items having indices in respect to difficulty and discrimination indices with their distractor efficiency. Very few authors keep penalty (Negative marks) for each incorrect answer. This penalty is an unbiased estimate of what a student can get when answering randomly if there is no penalty. The negative marking contribution is crucial for this type of test,[4] thus decides ranks.

The present study was aimed to evaluate the Difficulty Index, Discrimination Index and Distractor Efficiency of the items present in the question bank (Memory based) of postgraduate medical entrance examinations of India. Thus, objective of the present study is framing MCQ bank by exposing with penalty during routine assessment of MBBS students.

MATERIALS AND METHODS

Place of Study

Department of Physiology, Government Medical College Haldwani, Nainital, Uttarakhand.

Study Design

Observational cross-sectional study.

Study Subjects
1st year MBBS Students of academic session 2017.

Data Collection Method
After 47 hours of dedicated teaching of Physiology of Central Nervous System. A unit test was conducted (routine assessment) based on MCQs.

50 MCQs items were selected from standard question bank (Memory based) of Physiology usually followed by postgraduate aspirants. To avoid cheating, question papers were prepared in 4 sets with random sequence of questions. Adjacent students were provided different sets of question paper.

All MCQs are items with stem (Starting part of the question) and a set of 4 possible responses, of which there are one key (Correct response) and three distractors (Incorrect response).[5,6] The students have to select just one response or none.

Marks Allotment are as Follows
- Correct response= 1 mark.
- Incorrect response= negative 1/4th of a mark (-0.25).
- None response = 0 marks.

Total 97 out of 100 students participated in 60 minutes test. The result of student's performance in these MCQ based unit test was entered in MS Excel 2007 and arranged in descending order. Two groups were formed (The size of these groups varies according to the literature, but it is usually around 30% of the total number of students. The most frequent size in literature is 27%[4,7]).

Group I- The top 27% scorers (28 students), who obtained the correct responses were considered as Higher ability (H).
Group II- The bottom 27% scorers (28 students), who obtained the correct responses were considered as Lower ability (L).[8]

Each items were analysed for three parameters and calculated with the following formula.[9,10,11]

a) Difficulty Index (DIF-I): DIF-I define the percentage of students who answered the items correctly and ranges between 0 and 100%.[12]

$$DIF\text{- }I = [(H + L) / N] \times 100$$

b) Discrimination Index (DI): DI is ability of an item to distinguish between students of higher and lower capacities and ranges 0 to 1. It can be < 0 (negative DI), which means students of lower ability answer more correctly than students of higher ability. It indicates that students of lower ability respond the answer without any real understanding, just by guess; while a good student suspicious of an easy question, takes harder path to solve and ends up less successful.[5] Such situations are undesirable.

$$DI = 2 \times [(H - L) / N]$$

a) Distractor Efficiency (DE%) Analysis: An item has one correct (key) and three incorrect (distractor) options. A non-functional distractor (NFD) was defined as an option with a response frequency of < 5%, while the Functional Distractor (FD) are those selected by 5% or more participants.[11,13]

No. of NFDs	DE (%)
3	0
2	33.3
1	66.6
0	100

Table 1. Distractor Efficiency (DE%)[11,14]

RESULTS
Assessment of 50 items based on various indices amongst 97 students

DIF-I (%) Cut-Off Points	Items (N= 50)	Percentage	Interpretation
< 30	14	28%	Difficult
31-40	10	20%	Good
41-60	13	26%	Excellent
≥ 61	13	26%	Easy
Discrimination Index (DI)			
< 0	3	6%	Undesirable
< 0.15	13	26%	Poor
0.15-0.24	5	10%	Good
≥ 0.25	29	58%	Excellent

Table 2. Distribution of Items in Relation to DIF-I and DI

- Higher the DIF-I lower is the difficulty and lower the DIF-I higher is the difficulty. Greater the value of DI, items are more able to discriminate between students of higher and lower capacities. The Difficulty Index and Discrimination Index are reciprocally related.[14]

Parameter	Mean	Standard Deviation (SD)
Difficulty Index (DIF I%)	42.79	23.57
Discrimination Index (DI)	0.28	0.22
Distractor Efficiency (DE%)	59.73	27.72

Table 3. Mean and Standard Deviations (SD) for DIF-I, DI and DE% Respectively

Mean ± SD of DIF-I % is 42.79 ± 23.57, DI is 0.28 ± 0.22 and DE % 59.73 ± 27.72.

Parameter	Values
No. of items	50
No. of total distractor	150
Functional distractor	90 (60%)
Non-functional distractor	60 (40%)

Table 4. Distractor Analysis

No. of Items with NFD	Frequency	Percentage	Cumulative
0 NFD	10	20 %	20%
1 NFD	23	46 %	66%
2 NFD	14	28%	94%
3 NFD	3	6%	100%

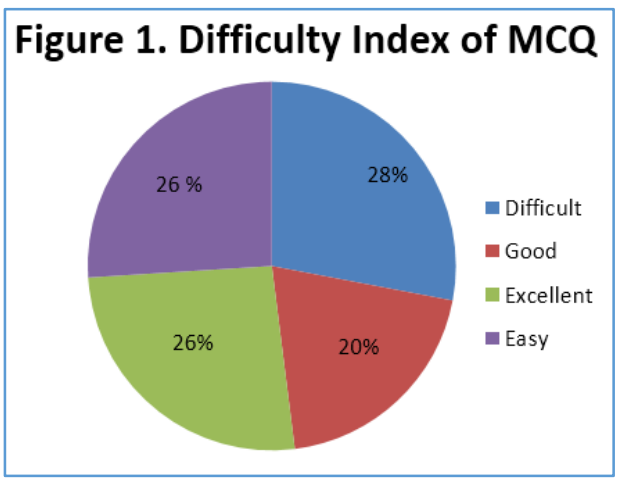
Table 5. Frequency distribution of Non-Functional Distractors (NFD) according to Selection

Out of 150 distractors, 60 (40%) NFD were present in 40 items (Table 5 and 6). 10 (20%) items had no NFDs, whereas

23 (46%), 14 (28%) and 3 (6%) items contained 1, 2 and 3 NFD respectively (Table 6).

DISCUSSION

Properly designed MCQs are one of the best assessment tools that quickly assess any level of cognition according to Bloom's taxonomy[15] and construction of appropriate test items question bank in form of quality MCQs is a big challenge. Developing an appropriate assessment strategy is a key part of curriculum development. Item analysis provides one such method of analysing observation and interpretation of the knowledge gained by the students.[16]



- Figure 1 shows that out of 50 questions 28% (14) are difficult, 20% (10) good, 26% (13) excellent and 26% (13) easy.
- So, the distribution of questions according to the difficulty level are ideal. It means questions for Doctor of Medicine are appropriate proportion of difficult and easy questions (i.e. neither too difficult nor too easy for the student).

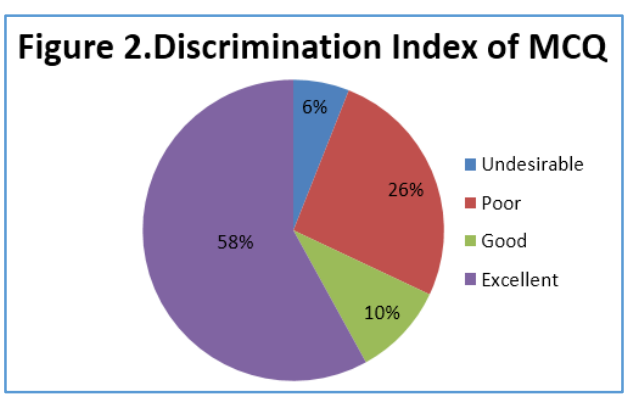


Figure 2 shows that out of 50 questions, 68% (34) questions differentiate very well between students of higher capacities from lower capacities. A question paper having >60% of DI can be considered as ideal for any examination, because the main objective of examination is to assess the knowledge and improve the skill of the students.

DIF-I in our study was 42.79 ± 23.57. Karelia et al (2013) showed a range of mean ± SD between 47.17 ± 19.77 and 58.08 ± 19.33 in their study.[17] Gajjar S (2014) showed mean and standard deviations (SD) for DIF-I (%) of 39.4 ± 21.4%.

Mean DI in present study was 0.28 ± 0.22, which was excellent. Gajjar S (2014) reported the items having DI > 0.25 have excellent Discrimination Index and Namdeo et al,[18] reported it as good.

Designing of plausible distractor and reducing the NFDs is important aspect for framing good quality MCQs. Presence or absence of NFD in an item affect the power. More NFD in an item easy and conversely item with more functioning distractors makes item difficult.[18] In our study among 150 distractors, 90 (60%) FDs and 60 (40%) NFDs were present. 10 (20%) items had no NFDs, whereas 23 (46%), 14 (28%) and 3 (6%) items contained 1, 2 and 3 NFD respectively (Table 6). In another study done by Namdeo et al (2016) showed 12%, 32%, 40% and 16% items had no, one, two and three NFDs respectively. This difference in our study might be due to penalty for incorrect response.

CONCLUSION

Item analysis of MCQs of previous postgraduate entrance examinations demonstrated Difficulty Index is neither too high nor too low. Discrimination Index is high, and Distractor Efficiency is good. Hence, we proposed that for formation of appropriate test items, question bank in the form of quality MCQs with negative marking should be prepared and their post validation, data to be collected.

ACKNOWLEDGEMENT

The statistical preparation was done with the help of department of Community Medicine, Govt. Medical College, Haldwani.

REFERENCES

[1] Hubbard JP, Clemans WV. Multiple-choice examinations in medicine: a guide for examiner and examinee. Lea & Febiger Libraries Australia 1961: p. 186.
[2] De Champlain AF, Melnick D, Scoles P, et al. Assessing medical students' clinical sciences knowledge in France: a collaboration between the NBME and a consortium of French medical schools. Acad Med 2003;78(5):509-17.
[3] Christian DS, Prajapati AC, Rana BM, et al. Evaluation of multiple choice questions using item analysis tool: a study from a medical institute of Ahmedabad, Gujarat. Int J Community Med Public Health 2017;4(6):1876-81.
[4] Amo-Salas M, Del Mar Arroyo-Jimenez M, Bustos-Escribano D, et al. New indices for refining multiple choice questions. Article ID 240263, Journal of Probability and Statistics 2014;2014: p. 1-8.
[5] Matlock-Hetzel S. Basic concept in item and test analysis, presented at annual meeting of the Southwest Educational Research Association, Austin: 1997: p. 1-27.
[6] Eaves S, Erford B. The Gale group: the purpose of item analysis, item difficulty, Discrimination Index, characteristic curve. 2007 edn. Online Source: www.education.com/reference/ article/itemanalysis/. Accessed on 21 April 2017.
[7] Tristr´an LA. Model for computer-aided item analysis. Foro Nacional de Evaluacion Educativa, Mexico: CENEVAL 1995: p. 45-68.

[8] Gajjar S, Sharma R, Kumar P, et al. Item and test analysis to identify quality multiple choice questions (MCQS) from an assessment of medical students of Ahmedabad, Gujarat. Indian J Community Med 2014;39(1):17-20.
[9] Singh T, Gupta P, Singh D. Principles of medical education. Test and item analysis. 3rd edn. New Delhi: Jaypee Brothers Medical Publishers (P) Ltd., 2009: p. 70-7.
[10] Scantron Guides - item analysis, adapted from Michigan State University website and Barbara gross devils tools for teaching. [Last cited on 2013 Apr 13]. www.freepdfdb.com/pdf/item-analysis-scantron.
[11] Hingarjo MR, Jaleel F. Analysis of one-best MCQs: the Difficulty Index, Discrimination Index and distractor efficiency. J Pak Med Assoc 2012;62(2):142-7.
[12] Sim SM, Rasiah RI. Relationship between item difficulty and discrimination indices in true/false type multiple choice questions of a para-clinical multidisciplinary paper. Ann Acad Med Singapore 2006;35(2):67-71.
[13] Tarrant M, Ware J, Mohammed AM. An assessment of functioning and non-functioning distractors in multiple- choice questions: a descriptive analysis. BMC Med Educ 2009;9:40.
[14] Mehta G, Mokhasi V. Item analysis of multiple choice questions - an assessment tool. Int J Health Sci Res 2014;4(7):197-202.
[15] Bloom S, Hastings T, Modays J. Handbook of formative and summative evaluation of students learning. New York: McGraw-Hill 1971: p. 103.
[16] Pellegrino J, Chudowsky N, Glaser R, (Book) Knowing what students know: the science and design of educational assessment. Washington, DC: National Academic Press 2001: p. 103-8.
[17] Karelia BN, Pillai A, Vagada BN. The levels of difficulty and discrimination indices and relationship between them in four response type multiple choice questions of pharmacology summative tests of year II MBBS students. IEJSME 2013;6:41-6.
[18] Namdeo SK, Sahoo B. Item analysis of multiple choice questions from an assessment of medical students in Bhubaneswar, India. Int J Res Med Sci 2016;4(5):1716-9.