STUDY ON ITEM ANALYSIS OF MULTIPLE – CHOICE QUESTIONS AMONGST THE UNDERGRADUATE DENTAL STUDENTS

Chhiring Palmu Lama,[1] Rabita Kharbuja,[1] Deepika Karki,[2] Shaligram Dhungel[1]

[1]Department of Human Anatomy, [2]Intern, Nepal Medical College Teaching Hospital, Attarkhel, Gokarneshwor-8, Kathmandu, Nepal

ABSTRACT

Assessments are powerful tools for bringing on the achievement of educational goals and one of the tools for bringing it rightly is "item analysis." Multiple - Choice Questions (MCQs) are one of the most common methods of assessing the knowledge capabilities of undergraduate and postgraduate medical students. The main objective of this study was to analyze the quality of MCQs and to assess the relationships of items having good difficulty and discrimination indices with their distracter efficiency. After getting an ethical approval from Institutional Review Committee of Nepal Medical College Teaching Hospital a cross – sectional study was conducted over a period of one year from January 2022 to December 2022 in the Department of Human Anatomy of Nepal Medical College. The difficulty index (DIF I), discrimination index (DI) and distracter efficiency (DE) were calculated and analyzed. The results obtained revealed that the mean and standard deviation of DIF I, DI and DE were 56.75±22.6, 0.3±0.2 and 43.24±25.66 respectively. The items analyzed were neither too easy and nor too difficult (DIF I = 62.8%). The majority of items fulfilled the criteria of acceptable difficulty and good discrimination index and the maximum number of items were found to have 2 NFDs (DE = 33.33 %). The present study was hence undertaken as it assists to evaluate the MCQ items to discern its effectiveness in assessing the knowledge of students as well as it plays a vital role in developing a question bank having valid MCQs.

KEYWORDS

Analysis, assessments, item, multiple choice questions

***Received on**: May 31, 2023*
***Accepted for publication**: August 02, 2023*

CORRESPONDING AUTHOR

Dr. Chhiring Palmu Lama
Assistant Professor
Department of Human Anatomy
Nepal Medical College and Teaching Hospital
Attarkhel, Gokarneshwor-8, Kathmandu, Nepal
Email: drcpalmu@yahoo.com
Orcid No: https://orcid.org/0000-0003-2560-8816
DOI: https://doi.org/10.3126/nmcj.v25i4.60876

INTRODUCTION

Assessments are powerful tools for bringing on the achievement of educational goals especially if conducted rightly. One of the tools for bringing it rightly is 'item analysis'.[1] The quality medical care depends upon the development of knowledgeable, skilled and competent medical personnel. Any assessment whether formative or summative has intense effect on learning and is an important variable in directing the learners in a meticulous way.[2] Multiple - Choice Questions (MCQs) are one of the most common methods of assessing the knowledge capabilities of undergraduate and postgraduate medical students.[3] Single correct response type MCQ is an efficient tool for evaluation.[4] MCQs were introduced into medical examination in 1950 as a reliable method of testing knowledge to replace traditional long essay questions.[5] MCQs were also found to be superior to the modified essay questions in assessing higher order skills.[6] Properly constructed MCQs can assess higher cognitive processing like interpretation, analysis and problem solving of Bloom's taxonomy instead of just recall of facts.[7] There is hardly any subject that cannot use MCQs. It is also used by examining bodies who conduct entrance examinations.[8]

The medical education across the world consists of initial assessment of the learner's need, monitoring the teaching-learning activities, certification of the competence to award a degree and practice medicine in context to the need of the society.[9] A good item can assess cognitive, affective as well as psychomotor domain and is preferred over other methods for its objectivity in assessment, comparability in different settings, wide coverage of subject and minimization of accessor's bias. Item analysis evaluates questions on three parameters. The difficulty of the questions that were asked can be analyzed by judging the difficulty index (DIF I). The discrimination index (DI) measures the ability of the item to discriminate good students from others. The distracter efficiency (DE) gives the idea of quality of distracters compared with the correct response.[10] An ideal item should have a DIF I between 30–70%, DI >0.2 and a DE 100%.[11] The periodic assessment of item analysis on different batches will enable the teachers to have a pool of 'good question banks' which also helps to get feedback on valid MCQ construction.[12] The present study was hence undertaken as its crucial to evaluate the MCQ items to discern its effectiveness in assessing the knowledge of students as well as it plays a vital role in curriculum development by creating an appropriate assessment strategy.

MATERIALS AND METHODS

The study was carried out after getting an ethical approval from Institutional Review Committee of Nepal Medical College Teaching Hospital, Gokarneshwor-8, Kathmandu, Nepal (Ref. No: 051- 078/079). A cross-sectional study was conducted over a period of one year from January to December 2022 in the Department of Human Anatomy of Nepal Medical College Teaching Hospital, Gokarneshwor-8, Kathmandu, Nepal. The study included 215 MCQs and 645 distracters from the internal assessment of anatomy of the undergraduate dental students. The study included anatomy MCQs asked in formative and sessional examinations and the MCQs which were not attempted by any students were excluded from the study. Each MCQ comprised a stem and four responses. A correct response to an item was awarded one mark and the wrong one zero. There was no negative marking. After the evaluation, the marks obtained by the students were arranged in descending order. The upper 1/3rd of the marks obtained were considered as high achievers and lower 1/3rd as low achievers. Marks obtained by middle 1/3rd were discarded. All MCQs included in this study were separately subjected to item analysis. Each item was analyzed for following three parameters.

1. Difficulty Index (DIF I) - It is the percentage of students in high or low achievers group who answered the item correctly. It ranges between 0 - 100 %. It was calculated by using the formula DIF I = H+L × 100/N.

   Where H = Number of students who answered the item correctly in high achieving group, L = Number of students who answered the item correctly in low achieving group, N = Total number of students in two groups including non – responders

2. Discrimination Index (DI) - is the ability of an item to differentiate between students of higher and lower abilities and ranges between 0 and 1. It was calculated by using the formula DI = 2× (H- L/N)

3. Distracter Efficiency (DE) - is determined for each item on the basis of number of non- functional distracters (NFDs) ie, option selected by <5 % of students[3]

It can be depicted as,

Items with 0 NFD = 100 % DE

Items with 1 NFD = 66.6 % DE

Items with 2 NFD = 33.33 % DE

Items with 3 NFD = 0 % DE [12,19,28]

Interpretations were done as follows:

1. Items with DIF I between 30-70% are acceptable, over 70% are very easy and below 30% are classified as difficult.
2. Items with DI between 0.21-0.35 are good, more than 0.35 are excellent and below 0.2 are poor.
3. Any of the distracters in the item which has not attracted even 5% of the total response is said to be non-functional distracters.[9]

The above calculated values were assessed and the descriptive statistical analysis including mean and standard deviation of DIF I, DI and DE were computed and analyzed by using a statistical tool as SPSS-16 and the findings were illustrated in a tabular form.

RESULTS

Total 215 MCQs and 645 distracters were evaluated in the study. The mean and standard deviation of DIF I, DI and DE were calculated and were recorded as 56.75 ± 22.6, 0.3 ± 0.2 and 43.24 ± 25.66 respectively. The values obtained were illustrated in Table 1.

The present study revealed that out of 215 items, 22 items (10.2%) were found to be of difficulty level (DIF I <30%) while 58 items (27%) were found to be very easy (DIF I >70 %) and remaining 135 items (62.8%) were items within an acceptable range (30-70 %) as illustrated in Table 2.

Table 1: Mean of DIF I, DI and DE

Parameters	Mean ± SD
DIF I	56.75 ± 22.6
DI	0.3 ± 0.2
DE	43.24 ± 25.66

While analyzing the items in relation to discrimination index, 93 items (43.3 %) were found to have poor discriminating power (DI≤0.2), 37 items (17.2%) were found to have good discriminating power (DI between 0.21-0.35) and remaining 85 items (39.5 %) were found to have excellent discriminating power (DI>0.35) as shown in Table 3.

Table 2: Analysis of items according to difficulty index

DIFI (%)	n	%	Item evaluation
<30	22	10.2	Difficult MCQ
30-70	135	62.8	Acceptable MCQ
>70	58	27	Very easy MCQ
Total	**215**	**100**	

On the basis of number of NFDs, items with DE 100% were 13 (6.0%), items with DE 66.6% were 70 (32.6%), items with DE 33.3% were 106 (49.3%) and items with DE 0% were 26 (12.1%) as shown in Table 4.

Table 3: Analysis of items according to discrimination index

DI	n	%	Discrimination power
≤ 0.2	93	43.3	Poor
0.21-0.35	37	17.2	Good
> 0.35	85	39.5	Excellent
Total	**215**	**100**	

Table 4: Analysis of items according to distracter efficiency

Items with number of NFDs		n	%
0	DE = 100.0 %	13	6.0
1	DE = 66.6 %	70	32.6
2	DE =33.3 %	106	49.3
3	DE = 0.0 %	26	12.1
Total		**215**	**100**

DISCUSSION

Various methods of assessments have been used for evaluation of medical students and single correct response type MCQ is an efficient tool for evaluation. The quality of MCQ is assessed by the analysis of each item. Though item analysis had been used in MCQ exams extensively, it can also be used to evaluate the quality of a standardized observed structured clinical examination (OSCE) checklist.[13] An item analysis can also be performed by using a Rasch model which could also provide a valuable information related to test reliability, item difficulty and examinee ability.[14]

As the DIF I differentiates the easy items from the acceptable and the difficult items, the present study will be of help to develop a question bank comprising MCQs.[26] The mean DIF I of the present study was 56.75±22.6, mean DI was 0.3±0.2 and the items having 1 NFD were 32.6 % which coincided with the findings of various other studies.[9,12,20,23,25] The present study showed that 62.8 % of items were of acceptable range as

far as difficulty index was concerned which was similar with the findings of other analysis.[15,16] The difficulty index analysis of present study recorded that 10.2% items were difficult and the maximum number of items were within the acceptable range (62.8%) which also concurred with the findings of other researchers.[18,19,21,22] If the MCQ is flawed, then it itself becomes distracting and assessment can be false. As this analysis also differentiates good MCQs from bad it also helps to revise, store, discard and to develop pools of MCQs. The current study recorded the excellent discrimination index (39.5%) as similar to the findings of other studies.[2,17] Similar study had also been found to be conducted in the department of histology in Ireland where the stems with images had been introduced to check the influence on item difficulty and to measure the discrimination.[16] An item analysis had also been found to be conducted to compare the performance between free and paying admission students in medical schools.[24] This analysis tends to be essential in assessing question's validity, reliability and its capacity in discriminating against the examinee's performance. The training and retraining of all faculty members are important to improve their skills in properly standardizing MCQs construction to overcome any assessment challenges.[27] A distracter analysis gives an opportunity to analyze the responses made by students on each item. If the students consistently fail to select certain distracters such items need to be modified. To conclude, the items analyzed in the present study were neither too easy and nor too difficult (DIF I =62.8%). The majority of items fulfilled the criteria of acceptable difficulty and good discrimination index and the maximum number of items were found to have 2 NFDs (DE =33.3%). The very easy items and the items having poor discriminating power will be revised and reconstructed. For easy items, discrimination may be poor as both high and low achievers can answer the item correctly. The items with good discriminating power tend to be moderately difficult items. More NFDs in an item increases DIF I. Hence, the findings of the current study will increase the awareness of this analyzing tool among the medical education providers in any field and the question bank thus created will be of great value in the respective department for future use.

Conflict of interest: None

Source of research fund: None

REFERENCES

1. Odukoya JA, Adekeye O, Igbinoba AO, Afolabi A. Item analysis of university- wide multiple choice objective examinations: the experience of a Nigerian private university. *Qual Quant* 2018; 52: 983-97.
2. Gajjar S, Sharma R, Kumar P, Rana M. Item and test analysis to identify quality multiple - choice questions from an assessment of medical students of Ahmedabad, Gujarat. *Indian J Com Med* 2014; 39: 17-20.
3. Kaur M, Singla S, Mahajan R. Item analysis of in use multiple - choice questions in pharmacology. *Int J App Basic Med Res* 2016; 6: 170-3.
4. Hingorjo MR, Jaleel F. Analysis of one – best MCQs: the difficulty index, discrimination index and distracter efficiency. *J Pak Med Assoc* 2012; 62: 142-7.
5. Premadasa IG. A reappraisal of the use of multiple - choice questions. *Med Teach* 1993; 15: 237-42.
6. Palmer EJ, Devitt PG. Assessment of higher order cognitive skills in undergraduate education: modified essay or multiple choice questions? Research paper. *BMC Med Educ* 2007; 7: 1-7
7. Khan MUZ, Aljarallah BM. Evaluation of modified essay questions and multiple choice questions as a tool for assessing the cognitive skills of undergraduate medical students. *Int J Health Sci* 2011; 5: 39-43.
8. Arora S. Item analysis. *Int J Nursing Sci Pract Res* 2018; 4: 1-5.
9. Bhat SK, Prasad KHL. Item analysis and optimizing multiple- choice questions for a viable question bank in ophthalmology: a cross-sectional study. *Indian J Opthalmol* 2021; 69: 343.
10. Ananthakrishnan N, Sethuraman KR, Kumar S. editors. Medical Education – Principles and Practice (2nd ed.). Alumni association of National teacher training centre, Jawaharlal Institute of Postgraduate Medical Education and Research: 2000.
11. Tarrant M, Ware J, Mohammed AM. An assessment of functioning and non – functioning distracters in multiple – choice questions: a descriptive analysis. *BMC Med Educ* 2009; 9: 40.
12. Adiga MNS, Acharya S, Holla R. Item analysis of multiple – choice questions in pharmacology in an Indian medical school. *J Health Allied Sci* 2021; 11: 130-5.
13. Donnon T, Lee M, Cairncross S. Using item analysis to assess objectively the quality of the Calgary – Cambridge OSCE checklist. *Canadian Med Edu J* 2011; 2: 1-9.

14. Yang SC, Tsou MY, Chen ET, Chan KH, Chang KY. Statistical item analysis of the examination in anesthesiology for medical students using the Rasch model. *J Chin Med Assoc* 2011; 74: 125-9.

15. Pande SR, Parate VR, Nikam AP, Agrekar SH. Correlation between difficulty and discrimination indices of MCQs in formative exam in physiology. *South East Asian J Med Edu* 2013; 7: 44-50.

16. Holland J, O'Sullivan R, Arnett R. Is a picture worth a thousand words: an analysis of the difficulty and discrimination parameters of illustrated vs. text- alone vignettes in histology multiple – choice questions. *BMC Med Educ* 2015; 15: 1-9.

17. Patil R, Palve SB, Vell K, Boratne AV. Evaluation of multiple – choice questions by item analysis in a medical college at Pondicherry, India. *Int J Com Med Pub Health* 2016; 3: 1612-6.

18. Namdeo SK, Sahoo B. Item analysis of multiple – choice questions from an assessment of medical students in Bhubeneswar, India. *Int J Res Med Sci* 2016; 4: 1716-9.

19. Ingale SA, Giri PA, Doibale MK. Study on item and test analysis of multiple – choice questions amongst undergraduate medical students. *Int J Com Med Pub Health* 2017; 4: 1562-5.

20. Kheyami D, Jaradat A, Al-Shibani T, Ali FA. Item analysis of multiple-choice questions at the Department of pediatrics, Arabian Gulf university, Manama, Bahrain. *Sultan Qaboos Univ Med J* 2018; 18: 1-11.

21. Angadi NV, Nagabhushana A, Hashilkar NK. Item analysis of multiple-choice questions of undergraduate pharmacology examinations in a medical college in Belgavi, Karnataka, India. *Int J Basic Clin Pharmacol* 2018; 7: 1-3.

22. Burud I, Nagandla K, Agarwal P. Impact of distracters in item analysis of multiple – choice questions. *Int J Res Med Sci* 2019; 7: 1-3.

23. Kundu S, Ughade JM, Sherke AR *et al*. Impact measurement on medical faculty for adhering to appropriate guidelines in framing effective multiple – choice questions for item analysis. *J Med Edu* 2020; 19: 1-15.

24. Rao C, HL Kishan Prasad, Sajitha K, Permi HS, Shetty J. Item analysis of multiple-choice questions: assessing an assessment tool in medical students. *Int J Educ Psychol Res* 2020; 2: 201-4.

25. Hosseini SM, Rahmati R, Sephery H, Habibi-Koolaee M. Comparison of the statics and item analysis of multiple-choice questions of medical and dental students test between free admission and tuition paying admission. *Res Square* 2020; 1-12.

26. Kumar B, Suneja M, Swee ML. Development and test - item analysis of a freely available 1900 item question bank for rheumatology trainees. *Cureus* 2021; 13: 1-60.

27. Elgadal AH, Mariod AA. Item analysis of multiple - choice questions: assessment tool for quality assurance measures. *Sudan J Med Sci* 2021; 16: 334-46.