

ITEM ANALYSIS OF MULTIPLE CHOICE QUESTIONS

¹ABDUL REHMAN, ²AYESHA ASLAM, ³SYED HAMMAD HASSAN

ABSTRACT

Any academic course or program is incomplete without an assessment plan. Of late, MCQs have been widely used for assessment of medical and dental undergraduate students. Often enough, poorly constructed MCQs are used that do not fulfil the requirement of an assessment. Therefore, it is necessary that quality of MCQs be frequently evaluated. A quality assessment of MCQs is known as "item analysis". Item analysis allows identification of good-quality MCQs based on their Difficulty Index (DIF I/P), Discrimination Index (DI) and Distractor efficiency (DE). The aim of this paper was to conduct an item analysis to evaluate the quality of MCQs being used to assess final year BDS students in the subject of Prosthodontics. An internal assessment for the subject of Prosthodontics was undertaken by fifty dental undergraduate students of final year in March 2018. The test consisted of 40 'one best type' MCQs, each with four options a – d. Post-validation of the test was done by item analysis. Data were analyzed using SPSS version 24. Majority (52.5%) MCQs were "too difficult" with a Diff Index less than 30%. About 67.5% MCQs had a poor discrimination index (DI < 0.2). Of the 120 distractors, 62 (51.6%) were non-functional and only 5 MCQs had no non-functional distractors. It can be concluded that the items administered in this test need careful revision to improve their quality before using them in any future assessments.

Key Words: Medical Education, Educational Assessment, Examination Questions, Discriminant Analysis

INTRODUCTION

Any academic course or program is incomplete without an assessment plan. An assessment plan refers to the proposed methods and time – line for assessment – related activities. Medical and dental health professions not only require sound knowledge and skill but also communication skills, analytical abilities, ethics etc.¹ This mandates that assessment methods in healthcare education be comprehensive, diverse and reliable so as to effectively assess the required attributes in addition to evaluating the basic knowledge and skill.

A number of different assessment tool are available to assess students' learning and performance. These include essay questions, short essay questions (SEQ), short answer question (SAQ), objective structured clinical examination (OSCE), extended matching items (EMI), multiple choice questions (MCQ), portfolios etc. Choice of an assessment method is based on the specific learning objectives to be assessed as well as the reliability, validity and feasibility of the assessment tool.²

Of late, MCQs have been widely used to for summative assessment of medical and dental undergraduate students.³

MCQs aid in assessing a large bulk of knowledge objectively in a short time.⁴ Moreover, if accurately constructed, they also help to assess the understanding and application of knowledge and problem – solving skills.⁵ An MCQ consists of a "stem" describing a scenario or background information followed by a "lead in" or the question statement.⁶ As the name describes, each MCQ has multiple options, usually four, including the answer to the asked question. Type A MCQs are described as "one – best type" where students have to choose the best possible answer from the given options.⁷

Advantages associated with MCQs such as validity, reliability and objectivity can only be appreciated if they are well-constructed. Often enough, poorly constructed MCQs are used for assessment that do not evaluate the required level of Bloom's taxonomy, are too easy or too difficult for the students, are ambiguous, can be easily 'guessed' or fail to differentiate between high ability and low ability students. Therefore, it is necessary that quality of MCQs be frequently assessed. Improvements can then be made according to the findings of the quality assessment and the items can be subsequently used in future tests.⁸

A quality assessment of MCQs is known as "item analysis". Item analysis allows identification of good-quality MCQs based on their Difficulty Index (DIF I/P), Discrimination Index (DI) and Distractor efficiency (DE).⁹ Difficulty Index helps identify whether an MCQ

¹ Dr Abdul Rehman, BDS, FCPS, Asst Professor, Prosthodontics, Armed Forces Institute of Dentistry, Rawalpindi

² Ayesha Aslam BDS, MSc., CHPE, Senior Lecturer, Armed Forces Institute of Dentistry National University of Medical Sciences (NUMS), Islamabad

³ Syed Hammad Hassan **For Correspondence:** BDS, FCPS, M.Sc. (Med Edu), Asst Professor, Prosthodontics Armed Forces Institute of Dentistry National University of Medical Sciences (NUMS), Islamabad

E-mail: drhammadhassan@gmail.com

Cell:03455348452

Received for Publication: June 10, 2018

Approved: June 28, 2018

is too easy or too difficult to attempt, Discrimination Index (DI) helps differentiate in high ability and low ability students while Distractor Efficiency (DE) highlights whether the distractors i.e. options other than the correct answer were effective or ineffective.⁷

MCQs are routinely being used to assess students in undergraduate dental programs. However, not enough emphasis is being laid on assuring and maintaining the quality of these test items. The aim of this paper was to conduct an item analysis to evaluate the quality of MCQs being used to assess final year BDS students in the subject of Prosthodontics. Knowledge thus gained will help in improving the MCQs and refining the quality of student assessment.

METHODOLOGY

An internal assessment for the subject of Prosthodontics was undertaken by fifty dental undergraduate students of final year in March 2018. The test consisted of 40 'one best type' MCQs, each with four options a – d. A total time of 60 minutes was allowed to solve the test. Post-validation of the test was done by item analysis. Data was analyzed using SPSS version 24. Difficulty index was calculated by the formula

$$P = (H+L/N)*100$$

where H is the number of students in higher ability group correctly answering the item and L is the number of students in lower ability group correctly answering the item and N is to total number of students.

Discrimination Index was calculated by the formula

$$DI = (H-L/N)*2$$

Distractor Efficiency, number of non-functional distractors (NFD) per item and number of items with non-function distractors were also calculated. NFDs were those options which were selected by <5% of the students. Distractor efficiency ranged from 0% – 100% as the number of NFDs in an item decreased. Mean $\pm$ standard deviation for all three parameters (Diff I, DI and DE) were calculated. Percentage of items falling in various categories of Diff I and DI were also calculated.

RESULTS

Fifty students took the test which comprised of 40 MCQs. A score of 1 was given for every correct answer. Test scores ranged from 10 – 28 with a mean test score of 20.18 ± 4.08 . Table 1 highlights the Mean $\pm$ SD of the three parameters of item analysis. Results of Diff I showed that majority (52.5%) MCQs were "too difficult" with a Diff I less than 30% (Figure 1). 67.5% of MCQs had a poor discrimination index with DI < 0.2 (Figure 2). Of the 120 distractors, 62 (51.6%) were non-functional (Table 2). Of the total 40 MCQs, only 5 (12.5%) had no non-functional distractors (Table 3).

DISCUSSION

The importance of assessment in medical and dental education cannot be overemphasized. A well – structured assessment allows the teacher to test higher levels of cognitive domain.¹⁰ It also aids in dis-

tinguishing between high achievers and low achievers. An important aspect, however, is the post-test analysis of the assessment. Tarrant et al.¹¹ proposed and

TABLE 1: MEAN VALUES OF VARIOUS INDICES USED IN ITEM ANALYSIS

Parameter	Mean	SD
Difficulty Index	25.75%	15.46
Discrimination Index	0.12	0.13
Distractor Efficiency	48.29%	31.07

TABLE 2: FREQUENCY OF NON-FUNCTIONAL DISTRACTORS IN TEST ITEMS

Distractor Analysis	
Total no. of items	40
Total no. of distractors	120
Functional Distractors	58 (48.3%)
Non-Functional Distractors	62 (51.6%)

TABLE 3: DISTRIBUTION OF TEST ITEMS ACCORDING TO THE FREQUENCY OF NON-FUNCTIONAL DISTRACTORS

No. of Non – Function- al Distractors	Number of Items	
	Frequency	Percentage
No NFD 0	5	12.5
1 NFD	15	37.5
2 NFD	13	32.5
3 NFD	7	17.5
Total	40	100

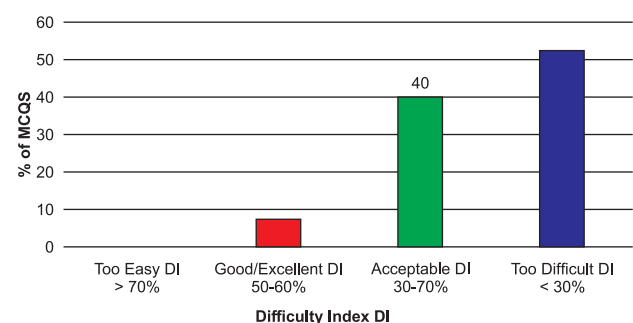


Fig 1: Percentage distribution of items (MCQs) according to their Difficulty Index emphasized the significance of reviewing the quality and performance of items after the administration of assessment test and improving the items for future use based on the results of review.

One-best type MCQs are an effective tool for assessment.⁷ However, their efficiency exclusively depends up on the quality of MCQ that can be assessed by post-test 'item analysis'.⁹ When it is a part of an assessment, each item must be analyzed for its difficulty, power

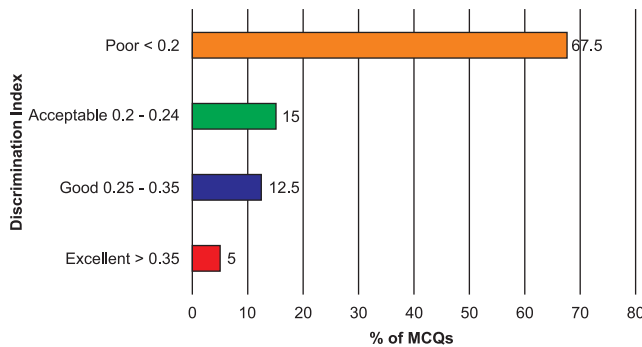


Fig 2: Percentage distribution of test items according to their Discrimination Index

to discriminate and for efficiency of its distractors i.e. alternative options. Mean Diff I in the present study was 25.75 ± 15.46 indicating that on average, items were too difficult to attempt for the students. This mandates that difficult items be thoroughly reviewed difficult items should be reviewed for any ambiguity, controversial alternatives, or even an incorrect key.⁹ Acceptable range for Diff I is 30 – 70%. About 40% of the items in the present study lied in the acceptable range while 52.5% were in the “too difficult” range. Another study reported 78% of its items in the acceptable range⁷ while Shete et al.⁸ reported 30% of their items in the acceptable range.

Discrimination index (DI) ranges from 0 – 1 and a greater value indicates increased ability of an item to differentiate between high achieving students and low achieving students. A value <0.2 indicates poor discriminative ability while a negative value shows a flawed item which was correctly answered only by low achieving students, probably by guessing. In the present study, mean DI was 0.12 ± 0.13 with 67.5% of the test items having a DI < 0.2. Similar results have been reported by Gajjar et al.⁹ Namdeo and Sahoo¹² reported a mean DI of 0.33 ± 0.23 which signifies good discriminative ability of test items. Hingorjo et al.⁷ reported a mean DI of 0.36 ± 0.17 with only 2 out of 50 items showing a negative DI. A negative DI may result from an incorrect key, poor or confusing structure of the item or poor preparation of students. In the present study, 6 out of 40 items had a negative DI. This may be attributed to poor preparation of students as revealed by a total score of (10-28)/40 in addition to flawed items.

Mean distractor efficiency (DE) in the present study was $48.29 \pm 31.07\%$. Of the 120 distractors, 51.6% were NFDs and majority of the items had 1 – 3 NFDs. Gajjar et al.⁹ reported a mean DE of 88.6 ± 18.6 that shows good efficiency of distractors while Hingorjo et al.⁷ reported a mean DE of 81.4%. Similar results were reported by Namdeo and Sahoo with 53.4% NFDs.¹² Presence of non-functional distractors increases the Diff I, making

an item easier to attempt. Subsequently, the DI of an item with greater number of NFDs will be poor.

The parameters of an item analysis help gauge the quality of administered assessments. Items with moderate difficulty, higher discrimination and functional distractors must be incorporated into tests to improve the test standard and quality. This would in turn help improve the overall test scores and properly discriminate among student of varying caliber. Teachers must be properly trained in assessment designing so that effective and meaningful test may be conducted.

CONCLUSION

Within the limitations of this study, it can be concluded that the items administered in this test need careful revision to improve their quality before using them in any future assessments. MCQ pools of all subjects must be critically analyzed for quality assessment. Post-test item analysis of all summative and formative assessments should be undertaken to improve the quality of healthcare education in the country.

REFERENCES

- 1 Tabish SA. Assessment Methods in Medical Education. *Int J Health Sci.* 2008;2(2):3-7.
- 2 Peck C. Principles of Sound Assessment Practice in Health Professions Education. *EC Psychology and Psychiatry.* 2017;5(5):150-7.
- 3 Mehta G, Mokhasi V. Item analysis of multiple choice questions - an assessment of the assessment tool. *Int J Health Sci Res.* 2014;4(7):197-202.
- 4 Patil PS, Dhobale MR, Mudiraj NR. Item analysis of MCQs' - Myths and realities when applying them as an assessment tool for medical students. *Int J Curr Res Rev.* 2016;8(13):12-6.
- 5 Al-Wardy NM. Assessment Methods in Undergraduate Medical Education. *Sultan Qaboos University Medical Journal.* 2010;10(2):203-9.
- 6 Mukherjee P, Lahiri SK. Analysis of Multiple Choice Questions (MCQs): Item and Test Statistics from an assessment in a medical college of Kolkata, West Bengal. *IOSR J Dent Med Sci.* 2015;14(12):47-52.
- 7 Hingorjo 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.
- 8 Shete AN, Kausar A, Lakhkar K, Khan ST. Item analysis: An evaluation of multiple choice questions in Physiology examination. *J Contemp Med Edu.* 2015;3(3):106-9.
- 9 Gajjar S, Sharma R, Kumar P, Rana M. 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.
- 10 Ghayur S, Iqbal M. A Simplified MCQs Interpretation Guide for Faculty. *Journal of the College of Physicians and Surgeons--Pakistan : JCPSP.* 2015;25(8):579-82.
- 11 M MT, 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.
- 12 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):1716-19.

CONTRIBUTIONS BY AUTHORS

- | | |
|------------------------------|--|
| 1 Abdul Rehman: | Study concept and design |
| 2 Ayesha Aslam: | Data collection, analysis and manuscript writing |
| 3 Syed Hammad Hassan: | Critical revision |