Original Research Paper

Volume - 13 | Issue - 10 | October - 2023 | PRINT ISSN No. 2249 - 555X | DOI : 10.36106/ijar

Pathology

USE OF ITEM ANALYSIS OF MULTIPLE CHOICE QUESTIONS IN GENERAL PATHOLOGY

Roopa Urs AN	Assistant Professor, Department of Pathology, Karwar Institute of Medical Sciences, Karwar, India.
Sarita Nair	Associate Professor, Karuna Medical College, Palakkad, Kerala.
PramodKumar GN	Professor & Head, Department of Forensic Medicine and Toxicology, Karwar Institute of Medical Sciences, Karwar, India.
Shankar M Bakkannavar*	Associate Professor, Department of Forensic Medicine and Toxicology, Kasturba Medical College, Manipal, Manipal Academy of Higher Education, Manipal – 576104, India. *Corresponding Author
Manjula MN	Tutor 'Statistician', Karwar Institute of Medical Sciences, Karwar, Karnataka, India.

ABSTRACT **Introduction:** Multiple choice questions (MCQs) can test a variety of skills like analysis and problem-solving. The quality of MCQs depends on the quality of distracters. Writing good MCQs requires training and practice. In entrance examinations, MCQs with good distracters plays important role in differentiating good and average candidates. The objective of this study was to demonstrate the use of item analysis as a tool to ascertain MCQ Validity in terms of difficulty index, discrimination index & distractor effectiveness. **Method:** 80 MCQs were prepared by faculty on the different topics of General Pathology. These questions were administered to 120 II MBBS students. After evaluation, the score was arranged in decreasing order. The whole list was divided into the first 30% (40) of students (high achievers) & the last 30% (40) (low achievers). The difficulty index (Dif I), discrimination index (DI), and distracter effectiveness (DE) were calculated using standard formulae. These MCQs and distracters were then classified into groups as per standard reference ranges of these parameters **Results:** The difficulty index of 54 (67.5%) items was in the acceptable range (Dif I= 30–70%), 6 (7.5%) items were too easy (>70%), and 20 (25%) items were difficult (<30%). Discrimination index of 13 (16.25%) items was excellent (>0.35), 37 (46.25%) items were good (0.25–0.35), and 24 (30%) items were poor (<0.2). A total of 80 items had 240 distractors. Amongst these, 20 (8.3%) were nonfunctional distracters, 220 (91.7%) were functional distracters. The discrimination index showed inconsequentially positive correlation with difficulty index (r = 0.217, P = 0.053). The maximum discrimination was observed in acceptable range (P = 30–70%). **Conclusion:** Items with acceptable range of difficulty index with excellent discrimination power and maximum DE (100%) are kept as viable question bank & utilized for further assessment test. Items with poor & negative discrimination are discarded.

KEYWORDS : Item analysis, multiple choice questions, Difficulty index, Discrimination index, distractor efficiency, Pathology

INTRODUCTION

Assessment refers to any formal or purported action to obtain information about the competence and performance of a student.[1] It is considered a major curricular component, at par with educational objectives and learning experiences. Objectivizing assessment is important in the education system, both for summative and formative purposes. One method of achieving this purpose is the widespread use of objective written items, Multiple Choice Questions (MCQs).[2] Multiple choice questions are being increasingly used in almost every kind of examination. The Competency-Based Medical curriculum (CBME -2019 onwards) by the National Medical Commission (NMC), has introduced MCQs (Multiple Choice Questions) for both summative and formative assessment in the curriculum for undergraduate (MBBS) students. MCQs shall be accorded weightage of not more than 20% of the total theory marks.[3] With frequent usage, the importance of item analysis for question banking has emerged and is used for creating a viable bank of MCQs. Item analysis is the process of analyzing the performance of a multiple-choice item after it has appeared in question paper.[4] The main purpose of item analysis is to determine whether the item is of appropriate level of difficulty or is it capable of discriminating between the knowledgeable and ill-informed students. Item analysis is also useful to get feedback on the functionality of alternatives to the correct responses in the item.[2] The aim of this study was to analyze the quality of items (MCQ's) having good difficulty and discrimination indices with their distractor efficiency for question banking & their effective usage in assessment.

MATERIAL & METHOD

This cross-sectional study was conducted in the Department of Pathology as a part of the internal assessment. Pre-validation of the paper was done by all the faculty members of the department before the assessment. A total of 120, second year MBBS students took MCQ's test comprising 80 questions with a single best response. There was no negative marking and the time allotted was 90 minutes. Each MCQ was having a single stem with four options comprising one correct answer and three distractors (incorrect answers). Each MCQ was assigned one mark. The maximum marks possible to score were 80 and the minimum was zero. Post validation of the paper was done by item analysis. The scores of all the students were arranged in order of merit in decreasing order. The upper one-third (n=40) of students were considered high achievers and the lower one-third (n=40) were low achievers. Paper with average scores, middle third (n=40) was excluded from the study. Each item was analyzed for difficulty index (Dif I), discrimination index (DI), and distractor efficiency (DE). The difficulty index or P value was determined using the formula Dif I = H + L/N ×100. Dif I represents the difficulty index, H represents the number of students answering the item correctly in the high achieving group, and L represents the number of students answering the item correctly in the low achieving group. N represents the total number of students in the two groups (including non-responders). The discrimination index was calculated by the formula DI=H-L/N x 2 where the symbols H, L and N represent the same values as mentioned before.

Items with a Difficulty Index (Dif I) between 30-70% are considered acceptable, those with values over 70% & below 30% are very easy & difficult respectively. Likewise, the items with a discrimination index between 0.25 to 0.35 are good, those with more than 0.35 are excellent and those with values below 0.2 are poor discriminators. Negative discrimination indicates a defective item or wrong key answer.[2,4]

An item contains a stem and four options including one correct (key) and three incorrect (distractor) alternatives. Nonfunctional distractor (NFD) in an item is the option, other than the key selected by <5% of students, and functional or effective distractor is the option selected by 5% or more students. DE ranges from 0% to 100%. If an item contains three or two or one or nil NFDs, then DE would be 0, 33.3%, 66.6%, and 100%, respectively.

Statistical Analysis

The data are reported as a percentage and mean plus or minus standard deviation (SD) of n items. The relationship between the Dif I and DI values for all items was determined using Pearson correlation analysis using SPSS 20.0

RESULT

A total of 120 students gave the test consisting of 80 MCQs. As seen in Table 1, the mean difficulty index was 40.10% while the mean discrimination index was 0.21. The distribution between difficulty indices (range 6.25–97.50) and discrimination indices (range -0.125–0.5) in all 80 MCQ items were analyzed.

A total of 80 items had 240 distractors. Amongst these, 20 (8.3%) were Non-functional distractors (NFDs), 220 (91.7%) were functional distractors (FD). Mean distractor efficiency was 91.65 ± 18.00 and distribution range from 0% to 100% [Table 2]. Out of the total of 80 items, difficulty indices of 7.5% (06) of MCQ items were easy (Dif I > 70%), about 25% (20) were difficult (< 30%) and the remaining 67.5% (54) of the items were within an acceptable range (30–70%) (Table 3). The discrimination indices (DI) for 80 items showed 30% (24) of the items with poor discrimination power (<0.2), and 16.25% (13) of the items exhibited excellent discrimination (>0.35). The remaining 46.25 % were acceptable and good (0.2 to 0.35) [Table 4]. The discrimination index showed inconsequentially positive correlation with difficulty index (r = 0.217, P = 0.053). The maximum discrimination was observed in the acceptable range (P = 30–70%).

DISCUSSION

Item analysis is a process that examines student responses to individual test items (questions) to assess the quality of those items and of the test.[5] It is especially valuable in improving items that will be used again in later tests, but it can also be used to eliminate ambiguous or misleading items in a single test administration. MCQs have limitations of psychomotor & affective domains not being assessed though it assesses the cognitive domain of learning with higher order thinking.[6,7]

Few authors termed Diff I as facility value indicated by the symbol 'P'. But it is a misnomer, as more is the Diff I, easier is the item, and vice versa.[1,2] In a study conducted by Rao C et al[5] on 120 students of pathology for 40 MCQs, mean Dif I 50.16± 16.15 was reported. Out of 40 items, 34 (85%) of the items were within accepted range (Dif I=30-70%), 2 (5%) were easy (Dif I=>70%), 4 (10%) items were difficult (Dif I=<30%). Item analysis done by Shahid R et al[8] on 336 4th MBBS students in Pathology for 50 MCQs, 20(40%) items was moderately difficult (25-75%), 30 (60%) Items were easy (>75%) with 0 items difficult (<25%). The study reported by Mahjabeen W et al[9] showed Mean Dif I of 58.74 ± 14.39. They conducted the test on 110 pathology students with 65 MCQs. 53 (81%) items were in an acceptable range, 11(17%) were too easy, and 1(2%) was difficult. In another study done by Kaur et al[10] on 150 students in Pharmacology for 50 MCQ mean Dif I was 59.18 ± 15.14. The 'P' value of 38 (76%) items were in the acceptable range (30-70%), 11(22%) items were easy (> 70%) and 1(2%) item was too difficult (<30%). The study conducted by Patil and Patil[11] on 100 MBBS students of medicine for 100 MCQs, showed a mean Dif I of 48.90 ± 13.72 was reported. In this study, the Dif I of 60 (47%) items was in the acceptable range (30–70%), 18 (18%) items were too easy (> 70%), and 22 (35%) items were too difficult (< 30%). Our study findings correspond with the study done by Patil and Patil having a mean Dif I of 40.10±17.90. The Dif I of 54 (67.5%)items was in acceptable range (30-70%), 6(7.5%) items were too easy (<70%) and 20 (25%) items were too difficult (<30%). Too difficult items (≤ 30%) can lead to deflated scores, while the easy items (> 70%) may result in inflated scores and a decline in motivation.[12,13] Items with high Dif I (>70%) should be placed either at the start of the test as "warm-up" questions to enhance the confidence of students or removed, similarly difficult items (<30%) should be either revised or removed altogether.[10,14] In our study 20 items were too difficult and removed from the list. Items that were too easy were 06 and these were revised and kept for subsequent use along with items within the acceptable range.

The discrimination Index of an item indicates its ability to differentiate between students of higher and lower abilities. It is apparent that a question that is either too difficult or too easy will have nil or poor DI.[12] In a study by Rao et al[5] out of 40 item, 24 (60%) item were excellent (DI >0.4), 4 (10%) items were good (DI= 0.3-0.39), 6 (15%) items were acceptable (DI=0.2-0.29) and 6 (15%) items were poor (DI <0-0.19). Mahjabeen W et al[9] study showed Mean DI of 0.35 ± 0.16 with 15 (23%), 5 (8%), and 11(17%) items that demonstrated good, acceptable, and poor discrimination respectively. Study reported by Kaur et al[10] showed mean DI of 0.37 ± 0.15 with 7 (14%) items were poor (DI < 0.2), 12 (24%) items had DI ≥ 0.20 and ≤0.35(good), and 31 (62%) items had DI > 0.35(excellent). In an item analysis study by Patil and Patil,[11] 24 item had DI < 0.2 (poor), 45 had good DI (≥ 0.20 and ≤ 0.35), and 31 had excellent DI (> 0.35). In a study done by Shahid R et al[8] (with 50 MCQ) showed mean DI of 0.27 ± 0.14 with 19 (38%) items showing poor DI, 17 (34%) items showing good DI, and 14(28%) items showing excellent DI. The present study findings with 80 MCQ were similar to this study and showed mean DI of 0.21 ± 0.14 with 24 (30%) items having poor, 37 (46%) items having good, and 13(16%) items having excellent DI. The mean DI in the present study (0.21 ± 0.14) was less than the acceptable cut-off points of 0.25 because 06 out of 80 items had negative (less than zero) DI. The study by Gajjar S et al[12] had a Mean DI of 0.14 ± 0.19 with 10 out of 50 negative DI. Another study by Hingorjo & Jaleel F[14] had 2 out of 50 items with negative DI. In our study, six items with negative DI were discarded because of the ambiguity since lower ability students answer questions correctly than those with the higher ability & tend to decrease the validity of the test.

Analysis of distractors is done to determine their usefulness in each item. Designing reasonable distractors & reducing NFDs is an important aspect of preparing MCQs.[12] The Study conducted by Kumar D et al,[15] for 90 items with a total of 270 distractors, 198 (73%) were functional distractors (FDs) and 72 (27%) were Non-functional distractors (NFDs). Of all items, two (2%) items had poor distractor effectiveness (DE) (0%), eight (9%) had moderate DE (33.3%), 50 (56%) had good DE (66.6%) and remaining 30 (33%) had excellent DE (100%). In a study of item analysis by Rao et al,[5] a total of 40 items with 120 distractors, mean DE was 89.99 ± 24.426. Out of 120 distractors, 6 (5%) were NFDs, 114 (95%) were FDs. The study by Kaur et al[10] reported 150 distractors for 50 MCQ. The mean DE was 83.98 ± 24.52 with 123 FDs and 27 NFDs. Items with no NFDs were 31. More NFDs in an item result in an increase in Diff I (easy item) and reduce DE, similarly item with more FDs decreases Diff I (difficult item) and increases DE. Our study showed Mean DE of 91.65± 18, with 220 (91.66%) FDs, 20 (8.33%) NFDs. So, Items with acceptable/ideal Diff I, the NFDs was modified and kept for future test assessment.

Table 1. Range, Mean & Standard Deviation Of Difficult, Discrimination Indices & Distractor Efficiency

	N	Range	Mean	Std. Deviation
Difficulty Index	80	6.25 - 97.50	40.10	17.90
Discrimination Index	80	-0.125 - 0.50	0.21	0.14
Distractor Efficiency	80	0 – 100	91.65	18.00

Table 2. Distractors And Categorization Of MCQS According To Distractor Efficiency (DE)

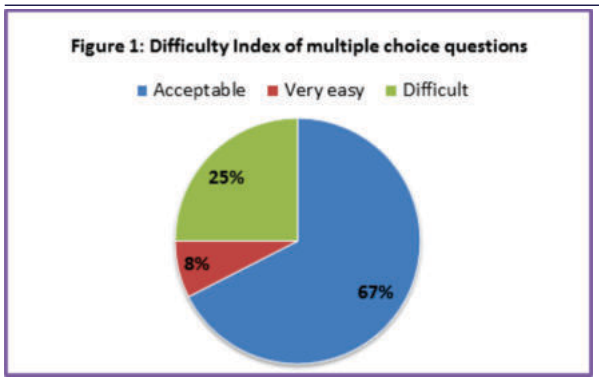
Parameter	Number (%)
Total MCQ	80
Total Distractors	240
Functional distractors	220 (91.66%)
Non functional distractors (NFD)	20 (8.33%)
Items with 0 NFD	63 (78.75%)
Items with 1 NFD	15 (18.75%)
Items with 2 NFD	1 (1.25%)
Items with 3 NFD	1 (1.255 %)
Mean DE	91.65 ± 18.00
Range	0% -100%

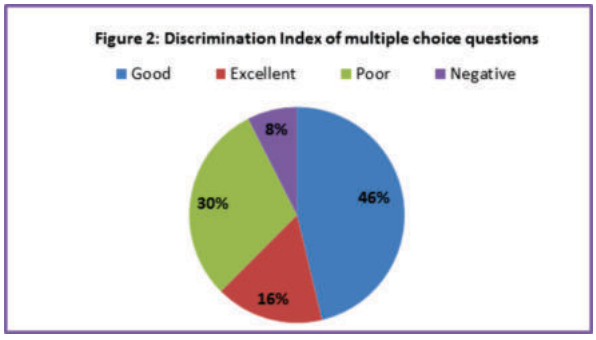
Table 3: Interpretation Of Difficulty Index Of MCQ Items (n=80)

Cut off point	Interpretation	Items (N=80) & Percentage	Action
30-70%	Acceptable	54 (67.5%)	Store
>70%	Very easy	06 (7.5%)	Revise
<30%	Difficult	20 (25%)	Revise /Discard

Table 4. Interpretation Of Discrimination Index Of MCQ Items (n=80)

Cut off points	Interpretation	Percentage	Action
0.25-0.35	Good	37 (46.25 %)	Store
>0.35	Excellent	13 (16.25 %)	Store
<0.2	Poor	24 (30 %)	Discard
	Negative	06 (7.5 %)	Discard





CONCLUSION

The item analysis is an effective procedure to evaluate the reliability and validity of an item. The difficulty index, discrimination index, distractor efficiency, and their interrelationship are used to analyse the items. Items with the acceptable range of difficulty index with excellent discrimination power and maximum DE (100%) will be kept as viable question banks & utilized for a further assessment test. Items with poor & negative discrimination are discarded. Items analyzed in the study were neither too easy nor too difficult (mean Dif I = 40.10±17.90) which is acceptable, but the overall DI was 0.21. Therefore, items were acceptably difficult but were poor at differentiating higher and lower ability students. DI was poor due to the 06 (7.5%) items with negative DI. Items with negative DI and NFDs will decrease the validity of the test & must be removed from the future assessment

REFERENCES

1. Singh T, Piyush G, Daljit S. Principles of Medical Education. 3rd ed. New Delhi: Jaypee Brothers; 2009. p. 70-7.
2. Ananthakrishna N. The item analysis, validation, and banking: In: Medical education principles and practice; by Anathakrishnan N, Sethukumaran KR, Kumar S, editors. ch. 20. Pondicherry, 2nd edn, India: Alumni Association of National Teacher Training centre, JIPMER; 2000. p.131-7.
3. https://www.nmc.org.in/wpontent/uploads/2020/01/Module_Competence_ based_ 02.09.2019.pdf assessed as on 6.8.2022 at 5pm
4. Ananthakrishna N, Ananthakrishnan S. The value of item analysis in classroom teaching. Indian J Med Edn 1991;30:22-25.
5. Rao C, Kishan Prasad HL, Sajitha K, Permi H, Shetty J. Item analysis of multiple-choice questions: Assessing an assessment tool in medical students. Int J Educ Psychol Res 2016;2:201-4.
6. Singh T, Anshu. Principles of Assessment in Medical Education. New Delhi: Jaypee Brothers Medical Publishers; 2012. p. 89.
7. Pande SS, Pande SR, Parate VR, Nikam AP, Agrekar SH. Correlation between difficulty & discrimination indices of MCQs in formative exam in physiology. South East Asian J Med Educ 2013;7:45-50.
8. Shahid R, Zeb S, Hayat U, Yasmeen S, Khalid MA. Umar M. Item Analysis of Pathology Assessment of 4thyear MBBS at Rawalpindi Medical University Pakistan. J Comm Med and Pub Health Rep 2021;2(5): https://doi.org/10.38207/jcmphr20210069
9. Mahjabeen W, Alam S, Hassan U, Zafar T, Butt R, Konain S, Rizvi M. Difficulty Index, Discrimination Index and Distractor Efficiency in Multiple Choice Questions. Ann. Pak. Inst. Med. Sci. 2017; 310-15
10. 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.
11. Patil VC, Patil HV. Item analysis of medicine multiple choice questions (MCQs) for under graduate (3rd year MBBS) students. Res J Pharm Biol Chem Sci 2015;6:1242-51
12. 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;39:17-20
13. Eaves S, Erford B. The Gale group: The purpose of item analysis, item difficulty, discrimination index, characteristic curve. Available from:www.education.com/ reference/article/itemanalysis. Assessed as on 8.8.2022 at 5pm
14. 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.
15. Kumar D, Jaipurkar R, Shekhar A, Sikri G, Srinivas V. Item analysis of multiple choice questions: A quality assurance test for an assessment tool. Med J Armed Forces India. 2021 Feb;77 (Suppl 1):S85-S89.