July - December 2023 Volume 32 Issue 2
PRINT ISSN: 2277-1867
ONLINE ISSN: 2277-8853

JOURNAL OF FORENSIC MEDICINE SCIENCE AND LAW

Official Publication of Medicolegal Association of Maharashtra

MULTISPECIALITY, MULTIDISCIPLINARY, NATIONAL
PEER REVIEWED, OPEN ACCESS, MLAM (SOCIETY) JOURNAL
Indexed with Scopus (Elsevier) & Index Copernicus (Poland)

Editorial Office Address
Department of Forensic Medicine & Toxicology, Third Floor, Library Building, Seth G S Medical College & KEM Hospital, Parel, Mumbai, Maharashtra, India. Pin-400 012. Email id: mlameditor@gmail.com Phone: 022-24107620 Mobile No. +91-9423016325.

JOURNAL OF FORENSIC MEDICINE SCIENCE AND LAW
(Official Publication of Medicolegal Association of Maharashtra)
Email.id: mlameditor@gmail.com

PRINT ISSN: 2277-1867
ONLINE ISSN: 2277-8853

Original Research Article

Post validation of Multiple Choice Questions in Forensic Medicine & Toxicology

Pramod Kumar GN[a], Roopa Urs AN[b], Jamshid P[c], Suraj S Shetty[d*]
[a]Professor & Head, FMT, Karwar Institute of Medical Sciences, Karwar, India. [b]Assistant Professor, Pathology, Chamarajanagar Institute of Medical Sciences, Chamarajanagar, Karnataka, India. [c]Assistant Professor, FMT, MES Medical College, Perinthalmanna, Kerala, India. [d]Additional Professor, FMT, K.S Hegde Medical Academy, Nitte (Deemed to be University), Mangalore, India.

Article Info

Received on: 23.12.2022
Accepted on: 25.10.2023

Key words
Assessment,
Difficulty Index,
Discrimination Index,
Distractor efficiency.

Abstract

Background: Item analysis examines student responses to individual test items (questions) to evaluate the quality of those items and the test as a whole. **Materials and Method:** The study was conducted in the Department of Forensic Medicine & Toxicology as a part of the internal assessment with 55 MCQs from Forensic Medicine subject. These questions were administered to 90 (batch of 115) students of fifth semester (second year MBBS students). Answer sheets were evaluated; the scores were then arranged in decreasing order. The whole list was then divided into the first 30% of students (high achievers) & the last 30% (low achievers). The difficulty index (Dif I), discrimination index (DI), and distractor effectiveness (DE) were calculated using standard formulae. These MCQs and distractors were classified as per standard reference ranges. **Results:** The difficulty index of 32 (58.18%) items was in the acceptable range (Dif I= 30–70%), 14 (25.45%) items were too easy (>70%), and 09 (16.36%) items were difficult (<30%). The discrimination index of 10 (18.18%) items were excellent (>0.35), 19 (34.55%) items were good (0.25–0.35), and 25 (45.45%) items were poor (<0.2). A total of 55 items had 165 distractors. Amongst these, 32 (19.75%) were non-functional distracters (NFD), 133 (80.60%) were functional distracters (FD). **Conclusions:** Post-validation of MCQs must be performed to filter MCQs of acceptable validity, which would increase their quality as assessment tools thereby making assessment more meaningful.

1. Introduction

Item analysis, also known as post-validation, is the process of analysing the performance of a multiple-choice question (MCQ) after it has appeared in a question paper/test.[1,2]

How to cite this article: Kumar P, Urs R, Jamshid P, Shetty S. Post validation of Multiple Choice Questions in Forensic Medicine & Toxicology. J For Med Sci Law 2023;32(2):51-55.

***Corresponding author:** Dr Suraj S Shetty, Additional Professor, FMT, K.S Hegde Medical Academy, Nitte (Deemed to be University), Mangalore, India. E-mail: drsurajshettyfmt@gmail.com (M): +919535625621.

Item analysis primarily creates a question bank of hundreds of questions with known difficulty levels and discrimination.[3]

The National Medical Commission's new curriculum makes objective assessment paramount in medical education. The Competency Based Medical curriculum (CBME -2019 onwards) has introduced MCQs (Multiple Choice Questions) for both summative and formative assessment for MBBS students in India. MCQs shall be granted weightage of not more than 20% of the total theory marks.[4] Framing of good MCQ is a time-consuming and challenging process. A well-constructed MCQ objectively measures knowledge, comprehension, application, analysis & evaluation.[5,6] Present study has been undertaken to evaluate MCQs or items and develop a good number of valid items by analysing with difficulty index (Dif I), discrimination index (DI), and distractor effectiveness (DE).

2. Materials and Methods:

This cross-sectional study was conducted in the Department of Forensic Medicine & Toxicology, Karwar Institute of Medical Sciences, Karwar, Karnataka 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 90 second year MBBS students took MCQ's test comprising 55 questions with a single best response. There was no negative marking, and the time allotted was 60 minutes. Each MCQ had a single stem with four options, including one correct answer and three distractors (incorrect answers). Each MCQ was assigned one mark. The maximum score possible was 55, and the minimum was zero.

Post-validation of the paper was done by item analysis. The scores of all the students were arranged in a decreasing order. The upper one-third (n=30) students were considered as high achievers and lower one-third (n=30) as low achievers. Paper with average scores, i.e., middle third (n=30), were excluded from the study. Each item was analysed for difficulty index (Dif I), discrimination index (DI), and distractor efficiency (DE). Difficulty index or P value was determined using 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, 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 before.

Items with Difficulty Index (Dif I) between 30-70% were 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.[2,3] An item contains a stem and four options, including one correct (key) and three incorrect (distractor) alternatives. Non-functional distractor (NFD) in an item is an option other than the key selected by <5% of students and a functional or practical distractor is the option chosen 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 (55) items. The relationship between the difficulty index and discrimination index values for all items was determined using Pearson's correlation coefficient and SPSS 16.0

3. Results:

A total of 90 students gave the test consisting of 55 MCQs.

Table 1: Range, Mean & standard deviation of Difficult, Discrimination Indices & Distractor efficiency (n=55)

Parameter	Range	Mean± Std Deviation
Difficulty index	13.3 – 91.7	52.20 ± 20.82
Discrimination index	0.03 – 0.73	0.22 ± 0.14
Distractor efficiency	33.3 – 100	83.57 ± 21.95

Table 2: Distractors and categorization of MCQs according to distractor efficiency

Parameter	Number (%)
Total MCQ	55
Total Distractors	165
Functional distractors	133 (80.60%)
Non functional distractors (NFD)	32 (19.75%)
Items with 0 NFD (DE 100%)	28
Items with 1 NFD (DE 66.6%)	22
Items with 2 NFD (DE 33.3%)	05
Mean DE	83.57± 21.95
Range	33.3 – 100

As seen in **Table 1**, the mean Dif I was 52.20 ± 20.82, while the mean DI was 0.22 ± 0.14. The

distribution between difficulty indices (range 13.3–91.7) and discrimination indices (range 0.03–0.73) in all 55 MCQ items was analysed. A total of 55 items had 165 distractors. Amongst these, 32 (19.75%) were Non-functional distractors (NFDs), and 133 (80.60%) were functional distractors (FD). The mean distractor efficiency was 83.57 ± 21.95, and the distribution ranged from 33.3% to 100% **(Table 2)**.

Figure 1: Difficulty index of multiple-choice questions

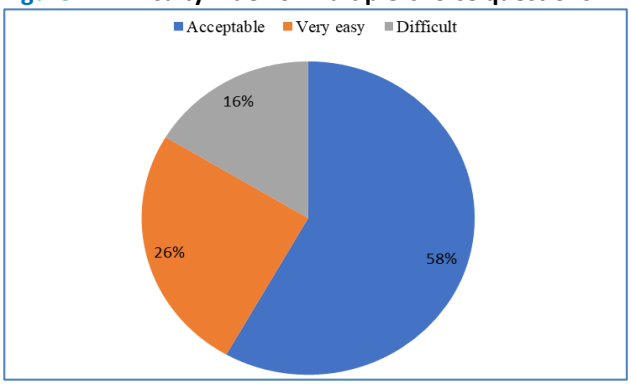
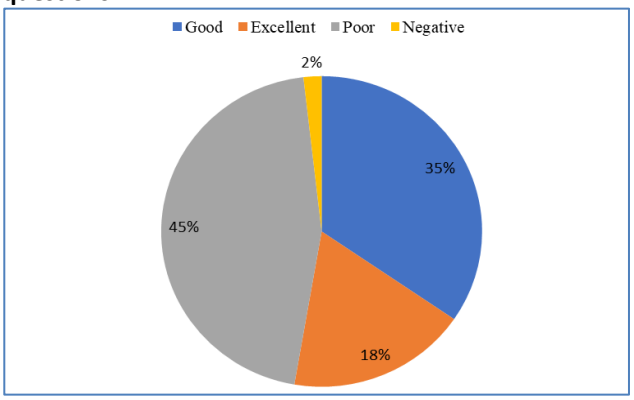


Figure 2: Discrimination index of multiple-choice questions



Out of a total of 55 items, difficulty indices of 25.45% (14) MCQ items were easy (Dif I > 70%), and about 16.36% (09) were difficult (< 30%). The remaining 58.18% (32) of the items were within an acceptable range (30–70%) **(Figure 1)**. The discrimination indices (DI) for 55 items showed 45.45% (25) of the items with poor discrimination power (<0.2), and 18.18% (10) of the items exhibited excellent discrimination (>0.35). The remaining 34.55 % (19) were acceptable and good (0.2 to 0.35) **(Figure2).**

4. Discussion:

Item analysis is particularly useful for improving items that will be used again in subsequent tests. It can also be used to eliminate ambiguous or misleading items in a single test administration. MCQs have the limitation of not assessing the psychomotor and affective domains, despite assessing the cognitive domain of learning with higher order thinking.[7,8] The parameter Difficulty Index is a misnomer. Few authors termed Diff I as a facility value indicated by the symbol 'P', as more is the Diff I, easier is the item, and vice versa.[2,5,9]

In a study conducted by Rao C et al. on 120 students of pathology for 40 MCQs, mean Dif I 50.16± 16.15 was reported.[10] Out of 40 items, 34 (85%) of the items were within the accepted range (Dif I=30-70%), 2 (5%) were easy (Dif I=>70%), 4 (10%) items were difficult (Dif I=<30%). The study reported by Mahjabeen W et al. showed a Mean Dif I of 58.74 ± 14.39.[11] 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 by Kaur et al. on 150 students in Pharmacology for 50 MCQ mean Dif I was 59.18± 15.14.[5] The Dif I 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 by Mehta G & Mokashi V with 50 MCQ for 100 students of anatomy showed Dif I as 63.06± 18.95.[12]

The Dif I of 31 (62%) items was in the acceptable range (30-70%), 16(32%) items were easy (>70%), and 3(6%) items were difficult (<30%). Our study findings correspond with the study done by Mehta & Mokashi V, having a mean Dif I of 52.20 ± 20.82. The Dif I of 32 (58.18%) items were in the acceptable range (30-70%), 14(25.45%) items were too easy (<70%), and 09 (16.36%) 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.[13,14] 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.[5,15] In our study, 09 items were too difficult and removed from the list. Items that were too easy[14] were, modified and kept for subsequent use, along with items within the acceptable range. Discrimination Index of an item indicates its ability to differentiate between students of higher and lower abilities & is used for selecting students as in an entrance examination.[2]

It is apparent that a question that is either too difficult or too easy will have nil or poor DI.[13] Generally, items of the middle level of difficulty are likely to have maximum discrimination. In a study by

Rao et al. out of 40 items, 24 (60%) items 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).[10] The study conducted by Mahjabeen W et al. showed Mean DI of 0.35 ± 0.16 with 15 (23%), 5 (8%), and 11(17%) items demonstrating good, acceptable, and poor discrimination, respectively.[11] The study reported by Kaur et al. showed mean DI of 0.37 ± 0.15 with 7 (14%) items being poor (DI < 0.2), 12 (24%) items being good (DI ≥ 0.20 and ≤0.35), and 31 (62%) items were excellent (DI > 0.35).[5] In a study done by Shahid R et al. (with 50 MCQs) showed a mean DI of 0.27 ± 0.14 with 19 (38%) items being poor, 17 (34%) items being good, and 14(28%) items being excellent.[16]

The present study findings with 55 MCQs were similar to this study and showed mean DI of 0.22 ± 0.14, with 25 (45%) items being poor, 19 (35%) items being good, and 10(18%) items being excellent in discrimination power. In our study, 1(2%) item had negative (less than zero) DI, and it was 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. Some other studies have shown negative DI in 10 (20%)13 and 2(4) % 15 MCQ items.

Analysis of distractors is done to determine their usefulness in each item. Designing plausible distractors & reducing NFDs is an important aspect of preparing MCQs.[13] In a study of item analysis by Rao et al. showed a total of 40 items with 120 distractors, mean DE was 89.99 ± 24.426.[10] Out of 120 distractors, 6 (5%) were NFDs, and 114 (95%) were FDs. Study by Kaur et al. reported 150 distractors for 50 MCQ.[5] The mean DE was 83.98 ± 24.52 with 123 FDs and 27 NFDs. In Mehta G and Mokashi V study 50 MCQs showed 150 distractors; 53 (35.33%) were found to be NFDs, and 28(18.66%) were FDs.[10] More NFDs in an item increase Diff I (easy item) and reduces DE. Similarly, item with more FDs decreases Diff I (difficult item) and increases DE.

Our study showed a Mean DE of 83.65± 21.95, with 133 (80.60%) FDs, and 32 (19.75%) NFDs. So, Items with acceptable Dif I and the NFDs were modified and kept for future test assessment.

5. Conclusion:

Items analysed in the study were neither too easy nor too difficult (mean Dif I = 52.20±20.82), which is acceptable, but the overall DI was 0.22. Therefore, items were acceptably difficult but were poor at differentiating higher and lower-ability students. DI was poor due to the one item with negative DI. Items with negative DI and NFDs will decrease the validity of the test & must be removed from the subsequent assessments.

Ethical Clearance: IEC approval is taken from the Institutional Ethical committee.

Contributor ship of Author: All authors equally contributed.

Conflict of interest: None to declare.

Source of funding: None to declare.

References:

1. Kishore KS, Pandey R. MCQ: Item analysis and banking. In PS Bhuiyan, NN Rege, AN Supe (Eds). The art of Teaching Medical students. 3rd edn: Elsevier; 2015. p 304-20.
2. Ananthakrishnan N. Item analysis, validation and banking of MCQs: In: Medical education principles and practice; by Ananthakrishnan N, Sethuraman KR, Kumar S, editors. Chapter. 20. Pondicherry, 2nd edn, India: Alumni Association of National Teacher Training centre, JIPMER; 2000. p.131-7.
3. Ananthakrishnan N, Ananthakrishnan S. The value of item analysis in classroom teaching. Indian J Med Edn. 1991; 30:22-5.
4. 4.NMC. [Cited 2022 Aug 06]. Available from: https://www.nmc.org.in/wpontent/uploads/2020/01/Module_Competence_based_02.09.2019.pdf
5. 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.
6. Kemp JE, Morrison GR, Ross SM. Developing evaluation instruments. In: Designing Effective Instruction. New York: MacMillan College Publishing Company; 1994. p. 180-213.
7. Singh T, Anshu. Principles of Assessment in Medical Education. New Delhi: Jaypee Brothers Medical Publishers; 2012. p. 89.
8. 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.
9. Singh T, Piyush G, Daljit S. Principles of Medical Education. 3rd ed. New Delhi: Jaypee Brothers; 2009. p. 70-7.
10. 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.
11. Mahjabeen W, Alam S, Hassan U, Zafar T, Butt R, Konain S et al. Difficulty index, discrimination index and distractor efficiency in multiple choice questions. Ann. Pak. Inst. Med. Sci. 2017;13(4):310-5.

12. Mehta G, Mokhasi V. Item analysis of multiple-choice questions – An assessment of the assessment tool. Int J Health Sci Res. 2014; 4:197-202.
13. 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
14. Eaves S, Erford B. The Gale group: The purpose of item analysis, item difficulty, discrimination index, characteristic curve. [Cited 2022 Aug 17] Available from: www.education.com/reference/article/itemanalysis
15. 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.
16. Shahid R, Zeb S, Hayat U, Yasmeen S, Khalid MA. Item Analysis of Pathology Assessment of 4th year MBBS at Rawalpindi Medical University Pakistan. J Comm Med and Pub Health Rep. 2021;2(5). [Cited 2022 Aug 17] Available from: https://doi.org/10.38207/jcmphr20210069