

ITEM ANALYSIS OF MULTIPLE-CHOICE QUESTIONS IN INTERNAL ASSESSMENT EXAMINATION QUESTION PAPERS OF PHASE I AND PHASE II MEDICAL UNDERGRADUATES.

B.Lavanya Devi¹, Rathnavel Kumaran Murugesan², N.Thilakavathi³, Usha Krishnan⁴

Received : 16/08/2025
Received in revised form : 04/10/2025
Accepted : 22/10/2025

Keywords:

Psychometric analysis, Multiple Choice Questions, Post validation, Difficulty Index, Discrimination Index, Distractor Efficiency.

Corresponding Author:

Dr. Usha Krishnan,

Email: ushakrishnan72@gmail.com

DOI: 10.47009/jamp.2025.7.6.46

Source of Support: Nil,

Conflict of Interest: None declared

Int J Acad Med Pharm
2025; 7 (6); 236-239



¹Associate Professor, Department of Biochemistry, Government Medical College, Omandurar, Tamil Nadu, India.

²Assistant Professor, Department of Physiology, Government Medical College, Thiruvallur, Tamil Nadu, India.

³Professor & Vice Principal, Department of Microbiology, Government Medical College, Thiruvallur, Tamil Nadu, India.

⁴Professor, Department of Microbiology, Madras Medical College, Chennai, Tamil Nadu, India.

ABSTRACT

Background: Innovation in preparing curriculum and adopting diverse assessment tool is evolving magnificently in education field. Multiple choice questions, being commonly used as formative and summative assessment tool, needs periodic analysis to assess its efficacy and validity. Psychometric analysis is a systematic methodological approach using statistical parameters such as Difficulty Index (DIF I), Discrimination Index (DI), and Distractor Efficiency (DE). The objective of this study was to do a structured stepwise analysis of the quality of MCQs and items appeared in Internal Assessment examination (IAE) of one subject from Phase 1 and phase 2 syllabus. **Materials and Methods:** In this descriptive study, IAE papers of 100 students in each of phase 1 and phase 2 of Biochemistry and Microbiology were included. Post validation of the item was done for the total of 100 MCQs and 400 items. **Result:** The results showed the DIF I with acceptable range of difficulty level (30 to 70%) was 36% in Phase 1 IAE and 62% in Phase 2 IAE. This was reflecting in the higher mean score of 7.26 out of 10 in Phase 1 IAE compared with mean score of 6.8 out of 10 in Microbiology IAE. This explicit that items in phase 2 IAE were more discriminatory with DI 42%(> 0.35 excellent) and 63% of items as functional distractors. In Phase 1 the beginning of MBBS curriculum, the DIF I was in easy range with good DE having 47% as functional distractors though DI was moderate. **Conclusion:** Psychometric analysis of MCQs is an essential tool to ensure valid and reliable assessments in medical education. Our study demonstrated that most items had acceptable difficulty and good discrimination, though several required revisions due to poor discrimination or non-functioning distractors. Regular psychometric evaluation can improve MCQ quality, enhance fairness, and support competency-based learning outcomes.

INTRODUCTION

Ensuring the quality of medical education is the prime prerequisite to produce competent health care personnel. Curriculum is the backbone for teachers which gives objectives, units of different activities and supportive reference materials.^[1] Learning and teaching are made more effective by new innovative teaching tools. In a competency-based curriculum, greater emphasis is given to the quality of assessment.^[2] An assessment tool with good reliability and validity can differentiate between satisfactory and unsatisfactory performers, it assesses the student understanding of medical concepts and its

application.^[3] Multiple choice questions (MCQs) are widely used due to their objectivity and efficiency. But constructing a good MCQs with less item writing flaws is labor intensive.^[4] Studies have shown that a well-constructed Multiple Choice Questions(MCQs) not only foster deeper learning but also provide a reliable measure of students' competencies.^[5] Psychometric analysis of MCQs is a systematic approach to evaluate the performance of individual test items and the overall test. The statistical tools to assess items are difficulty index, discrimination index, and distractor efficiency, which ensures the quality of the assessment.^[6,7]

By identifying poorly performing questions and optimizing test design, psychometric analysis enhances the fairness and effectiveness of the assessments. Moreover, the insights gained from psychometric analysis contribute to continuous improvement in curriculum design and student outcomes.^[8] Thus, integrating psychometric analysis into the evaluation of MCQs is indispensable for ensuring the alignment of assessments with educational objectives and for maintaining high standards in medical education.^[9]

However, only little light has been spotted on psychometric qualities of MCQs and its usage. But the current need on effective predictive validity of the every assessment tool is the backdrop to plan this post validation study with the objectives to process a simple step wise approach to perform the psychometric analysis of MCQs appeared in internal assessment exams (IAEs) conducted by medical Biochemistry and medical Microbiology.

Objectives

Objectives of this study is to understand and apply the psychometric qualities of Difficulty Index(DIF I), Discrimination Index (DI), and Distractor Efficiency (DE)of MCQs appeared in internal assessment exams (IAEs) conducted by medical Biochemistry and medical Microbiology.

MATERIALS AND METHODS

This descriptive, study was conducted over a period of two months, from April 2024 to May 2024 and evaluated five Microbiology IAEs comprising 50 MCQs and 3 Biochemistry IAEs comprising of 50 MCQs.

All MCQs were analyzed after the items had appeared in the question paper as post validation. The following are the simple steps to proceed with the psychometric analysis of each item.

Step -1: The test paper was arranged in rank order based on the student's performance from the highest score to the lowest score. The upper one third of the papers with high scores were named as the upper group. The Lower one third of the papers with low scores were named as the lower group. Only the upper and lower group were taken for the calculation. Step 2: For each item a frequency table was prepared by counting the number of students in the upper

group who selected each alternative. Similar count was done for lower group also.

Step 3: The psychometric parameters were calculated as given below.

A. Difficulty index (DIF I): It checks whether the difficulty level of each item is appropriate for the students. It is calculated using the formula

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

H= no. of correct responses in the high achieving group (upper group)

L= no. of correct responses in the low achieving group (lower group)

N= Total no. including non-responders in both the group

If the DIF I is less than 30%, the item is considered very difficult; if it is more than 70 %, it is considered easy. However, the DIF ranging from 30 % to 70 % was taken as an acceptable range. Higher the difficulty index lower is the difficulty of the question.

B. Discrimination Index (DI): It measures the ability of each item to discriminate between knowledgeable & ill-informed students. It is calculated by the differences in mark obtained for the correct responses between the higher achieving and the lower achieving group by the formula given below.

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

It has a value that ranges between 0 and 1. An item with DI of <0.15 is considered as poor, 0.15-0.25 is considered as marginal which can be improved to bring it to the good level i.e.>0.25-0.35. An item with>0.35 is considered as having excellent DI.

C. Distractor efficiency (DE): It is the measure of item functioning. When a distracter is chosen by >5% of participants, it is considered a functioning distracter (FD), and if chosen by <5% of participants is a non-functioning distractor (NFD).

RESULTS

A total of 100 students in each of phase 1 and phase 2 participated in the IAE of Biochemistry and Microbiology. Post validation of the item was done for the total of 50 items in each of the subject after the item has appeared in the IAE. An item contains a 'stem' which gives the background information to answer, the 'lead on' is the actual question and four options including one correct (key) and three incorrect (distractor) alternatives.

Table 1: Psychometric analysis of MCQs of Microbiology and Biochemistry IAEs.

Variables	Phase 2-Microbiology			Phase 1-Biochemistry	
	Range of Test score	2.8 to 10		4-10	
	Mean Test Score	6.82		7.86	
Psychometric parameter	Category	No. of MCQs in each category	Percentage (%)	No.of MCQs in each category	Percentage (%)
Difficulty Index	< 30% difficult	3	6%	2	4%
	30-70% acceptable	29	58%	17	34%
	>70% easy	18	36%	31	62%
Discrimination Index	< 0.15 Poor	14	28%	16	32%
	0.15-0.25 Marginal	6	12%	8	16%
	0.25-0.35 Good	9	18%	15	30%
	>0.35 Excellent	21	42%	11	22%

Distractor Efficiency	No of Distractor Assessed	150		150	
	Non Functional Distractor/Item	56	37%	79	53%
	Functional Distractor/Item	94	63%	71	47%
	3 NF Distractors	18	36%	11	22%
	2 NF Distractors	14	28%	15	30%
	1 NF Distractors	12	24%	14	28%
	0 NF Distractors	6	12%	10	20%

[Table 1] depicts the Psychometric analysis of MCQs. Base on DIF I percentage of distribution of MCQs as difficult: acceptable: easy range were 6%: 58%: 36% with the range of test score 2.8 to 10 (mean score - 6.82) in phase 2 Microbiology and were 4%: 34%: 62% with the range of test score 4 to 10 (mean score -7.86) in phase 1 Biochemistry. In this study items with negative discrimination index were nil. In Microbiology 40% and in Biochemistry 48% of MCQs had poor to marginal DI, also 60% and 58% of the items in respective subjects had a good to excellent DI.

A total of 300 distractors from 100 MCQs were analysed for the functionality of the distractors. In this study DE analysis expressed 47% and 63% of items were functional distractors whereas 53% and 37% of items were non-functioning distractors across Biochemistry and Microbiology IAE respectively.

DISCUSSION

Though MCQs remain the most commonly used assessment method in medical education constructing a quality MCQ is time-consuming and also difficult. Item analysis of MCQs in Competency Based Medical Education assessment is the process of collecting, summarizing and using information from students' responses to assess the quality of test items.

In the present study, items having <30% of Difficulty index that can be categorized as difficult, were found to be 4-6%. This is in contrast to a study evaluating item analysis of online IAE which reported that 15% of the items were found to be difficult.^[10] In this study based on the distribution of items as difficult, acceptable and easy category, Phase 2 Microbiology IAE had 58% of acceptable items which is contrast to study done by Kumar P et al with mean DIF I as $39.4 \pm 21.4\%$,^[11] and similar to another study done by Pande SS et al with mean DIF I as $52.53 \pm 20.59\%$.^[12] In Phase 1 Biochemistry IAE, the test score range was from 4 to 10 and had a high mean score 7.86 with more easy items (DIF I 62%). In the phase 2 Microbiology easy items had DIF I of 36% and the test score range was 2.8 to 10 (mean score - 6.82).

The psychometric analysis of IAE papers revealed that about 28-32% of the items had poor DI which is similar to a study analyzing the IAE of Phase I subject reporting 20% of poor DI.^[13] In the Microbiology and Biochemistry IAE papers the items with >0.25 of discrimination index having good discrimination were found to be 60% and 52% respectively. Any

discrimination index of 0.2 or higher is acceptable, and the test item would be able to differentiate between weak and good students. This is in contrast to a study analysing an assessment tool in medical students, showed that 75% had discrimination index of more than 0.2.^[14]

A distractor efficiency (DE) is the list of distractors that distract. The DE of a MCQs is calculated as 100%, 66%, 33% or 0%, if all the three distractors are distracting, two distractors chosen, one distractor chosen or all distractors not chosen respectively. Previous research has noted that the most difficult aspect for the item writers is to generate plausible distractors.^[15] A functional distractor is distractors that has been attempted by at least 5% or more of the students.^[16] The construction of MCQs with functional distractors is complex and might play an essential role in DI.

In this study we observed DE of 47% and 63% and non-functioning distractors accounting for 53% and 37% of questions across Biochemistry and Microbiology IAE respectively. This is similar to a study which reports only 52.2% of all distractors were found to be functioning effectively.^[16] An another study has reported that the percentage of items with three functioning distractors in most tests ranged from only 1.1 to 8.4% of all items which is in contrast to the present study in which 22-36% of the questions had three functional distractors.^[17] Items with more functioning distractors were found to be more difficult and more discriminating. About 28-30% of the items had two and one fourth of items had one functional distractors and 12-20% had zero functional distractors.

The present study underscores the significance of psychometric analysis as a practical tool to improve the quality of MCQs used in internal assessments. By identifying items with suboptimal difficulty, poor discrimination, and non-functional distractors, educators can refine question design to better evaluate students' competencies. This ensures fairer assessments, fosters deeper learning, and contributes to producing competent and confident medical graduates who are well-prepared for clinical practice. Way ahead, we had the plan to organize the Faculty development programs focusing on MCQ construction and distractor design which can further improve the quality of MCQs. To incorporate the psychometric analysis as a routine post-validation exercise to continuously enhance question quality.

CONCLUSION

Psychometric analysis of MCQs provides an objective and systematic method to evaluate the quality of assessment tools in medical education. The present study highlights that the majority of MCQs in both Biochemistry and Microbiology IAEs fell within the acceptable difficulty range and demonstrated good to excellent discrimination power. However, a considerable proportion of items had poor to marginal discrimination indices, indicating the need for revision and improvement of such items. Distractor efficiency analysis revealed that a significant number of distractors were non-functional, reinforcing the importance of constructing plausible distractors to enhance the validity of MCQs.

Overall, integrating psychometric evaluation into routine post-validation of assessments allows educators to refine question quality, improve reliability, and ensure alignment with competency-based educational objectives. Regular analysis and revision of MCQs will not only strengthen the assessment process but also promote deeper learning and better differentiation between high- and low-performing students, ultimately leading to the development of more competent healthcare professionals.

REFERENCES

- Sherratt Patrice. Principles of Curriculum Development, Implementation and Evaluation. Available from: https://www.academia.edu/Curriculum_Innovations. [Last accessed on 2019 Oct 12].
- Barriga FD, Barrón C. Innovation in education and curriculum reforms in Mexico. *Transnatl Curriculum Inquiry* 2012;9. Available from: <http://www.nitinat.library.ubc.ca/ojs/index.php/tci>. [Last accessed on 2019 Oct 12].
- Ebel RL. *Essentials of Educational Measurement*. 1st ed. New Jersey: Prentice-Hall; 1972.
- Palmer E, Devitt P. Assessment of higher order cognitive skills in undergraduate education: modified essay or multiple choice questions? *BMC Med Educ*. 2007;7:49.
- Schuwirth, L. W. T., & Van der Vleuten, C. P. M. (2011). Programmatic assessment: From assessment of learning to assessment for learning. *Medical Teacher*, 33(6), 478–485. <https://doi.org/10.3109/0142159X.2011.565828>.
- Halikar SS, Godbole V, Chaudhari S. Item Analysis to Assess Quality of MCQs. *Indian J Appl Res*. 2016; 28;6(3);123-125.
- Kehoe J. Basic item analysis for multiple-choice tests. *Practical Assessment, Research and Evaluation* 1995;4. Available from: <http://PAREonline.net/getvn.asp?v=4 and n=10>. [Last accessed on 2019 Oct 9].
- Karim Eldin M A Salih , Abubakar Jibo , Masoud Ishaq , Sameer Khan , Osama A Mohammed , Abdullah M AL-Shahrani , Mohammed Abbas , Psychometric analysis of multiple-choice questions in an innovative curriculum in Kingdom of Saudi Arabia. *J Family Med Prim Care*. 2020 Jul 30;9(7):3663–3668. doi: 10.4103/jfmpe.jfmpe_358_20.
- Mitra NK, Nagaraja HS, Ponnudurai G, Judson JP. The levels of difficulty and discrimination indices in type A multiple-choice questions of pre-clinical semester one multidisciplinary summative tests. *ASME* 2009;3:2-7.
- Bhattacharjee, Sharmistha; Mukherjee, Abhijit; Bhandari, Kallol; Rout, Arup Jyoti. Evaluation of Multiple-Choice Questions by Item Analysis, from an Online Internal Assessment of 6th Semester Medical Students in a Rural Medical College, West Bengal. *Indian Journal of Community Medicine* 47(1):p 92-95, Jan-Mar 2022. | DOI: 10.4103/ijcm.ijcm_1156_21).
- Kumar P, Sharma R, Rana M, Gajjar S. 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.
- Pande SS, 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 Educ*. 2013;7(1):45-50.
- Dharmendra Kumar, Raksha Jaipurkar, Atul Shekhar, Gaurav Sikri, V. Srinivas. Item analysis of multiple choice questions: A quality assurance test for an assessment tool. *Medical Journal Armed Forces India*, Volume 77, Supplement 1, 2021. Pages S85-S89. ISSN 0377-1237, <https://doi.org/10.1016/j.mjafi.2020.11.007>).
- 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)
- Rodriguez MC. Three options are optimal for multiple-choice items: A meta- analysis of 80 years of research. *Educ Meas Is Pract*. 2005;24(2):3-13.
- 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 Jul 7;9:40. doi: 10.1186/1472-6920-9-40. PMID: 19580681; PMCID: PMC2713226.
- Haladyna TM, Downing SM. How many options is enough for a multiple-choice test item? *Educ Psychol Meas*. 1993;53:999–1010.