



IMPACT OF EDUCATIONAL INTERVENTION ON ITEM DIFFICULTY, DISCRIMINATION, AND DISTRACTOR EFFECTIVENESS: A QUASI-EXPERIMENTAL STUDY IN MEDICAL EDUCATION

Syed Hyder Raza^{1*}, Muneeb-Ur-Rehman Niazi², Farheen Zehra³, Manal Iftikhar Niazi⁴, Rohab Asghar⁵

^{1*}Professor, Department of Pharmacology, Niazi Medical & Dental College, Sargodha, Pakistan.

²Assistant Professor, Department of Orthopedic, Niazi Medical & Dental College, Sargodha, Pakistan.

³Senior Registrar, Department of Medical Education, Niazi Medical & Dental College, Sargodha, Pakistan.

^{4,5}Niazi Medical & Dental College, Sargodha, Pakistan.

Corresponding Author: Prof. Dr. Syed Hyder Raza Naqvi

MBBS, MBA, CHPE, CHR, CACAP, M.phil (Pharmacology), Ph. D (Pharmacology)

ORCID ID: 0009-0000-2019-7734

Web of Science Researcher ID: OGN-1599-2025

Department of Pharmacology & Therapeutics

Niazi Medical & Dental College, Sargodha, Pakistan

Email: hyder.raza891@gmail.com

ABSTRACT

Background: Niazi Medical and Dental College (NMDC), Sargodha, has redefined the standards of medical pedagogy in the region since 2019. In keeping with this pursuit of excellence, the college utilizes MCQs as a primary modality for evaluating student competency, recognizing their critical role in valid and reliable written examinations.

Material & Methods: This quasi-experimental study was conducted at multiple clinical specialties, Niazi Medical and Dental College (NMDC), Sargodha, to evaluate the impact of faculty training on the quality of Multiple Choice Questions (MCQs). A total of 20 house officers from the four different specialties, divided 5 in each of four groups, with no formal psychometric training, were asked to develop 10 MCQs. Subsequently, a three-hour seminar on standard MCQ construction guidelines was conducted. Following this intervention, the original items were revised in accordance with established pedagogical standards and re-administered to the same cohort. Comparative psychometric analysis was performed using Microsoft Excel and SPSS version 22.0, focusing on: Difficulty Index (P), Discrimination Index (DI), and Distractor Effectiveness (DE).

Results: The educational intervention led to a statistically significant improvement in both student outcomes and the psychometric quality of the assessment items. Following the modification of MCQs in accordance with standard guidelines, the house officers' pass rate increased markedly from pre-seminar (62.5%) to post-seminar (87.5%), the Difficulty Factor improved from 0.42 to 0.55, and the Discrimination Index rose from 0.38 to 0.52.

Conclusion: Evidence suggests that the targeted workshop served as a critical catalyst for improving the assessment literacy of medical educators, resulting in a measurable enhancement in the quality of test construction.

Keywords: Medical Education, Educational Intervention, Student Competency.

INTRODUCTION

In the contemporary landscape of medical pedagogy, Multiple Choice Questions (MCQs) serve as the gold standard for objective written assessment.¹ Their utility lies in their capacity to evaluate diverse cognitive domains, ranging from fundamental recall to complex clinical comprehension.²

A standard MCQ consists of a "stem" (the problem statement), a "key" (the correct response), and a series of "distractors"—incorrect options designed to be plausible to students with incomplete mastery of the subject matter.^{3,4}

Institutional Context: NMDC, Sargodha

Since its establishment in 2019, **Niazi Medical and Dental College (NMDC)** has pioneered a rigorous approach to medical education in Punjab. A distinctive feature of the NMDC academic framework is the regulatory mandate requiring **100% MCQ-based theoretical evaluations**. This policy places a significant burden of reliability on the faculty, as these assessments constitute the



www.ajmrhs.com
eISSN: 2583-7761

Date of Received: 16-05-2026
Date Acceptance: 23-05-2026
Date of Publication: 22-06-2026

entirety of the summative data used for student certification and progression.⁵

Problem Statement & Rationale

Despite the high stakes of these examinations, a significant portion of junior faculty and demonstrators at NMDC began their academic roles with limited formal training in psychometrics or standardized assessment design.⁶ This "training gap" is critical; poorly constructed MCQs⁷—often characterized by structural flaws⁸ or non-functional distractors⁹—can compromise the validity of the entire examination process.^{6,10}

Study Objective

The **Department of Pharmacology and Therapeutics** recognized the urgent need for faculty development to align internal assessments with international standards. This study was designed to quantitatively evaluate the efficacy of a targeted educational intervention. By conducting a comparative psychometric analysis of items developed **before** and **after** a structured training seminar, the research seeks to measure the improvement in the quality of assessments generated by junior doctors at Niazi Medical and Dental College.

MATERIALS AND METHODS

Study Design and Ethical Considerations

A quasi-experimental (pre-post intervention) study was conducted in December 2025 at the **Niazi Medical and Dental College (NMDC), Sargodha**. Formal ethical clearance was obtained from the Institutional Ethical Review Committee (Ref. No. 123456789; dated 27th November 2025). All participants provided informed consent, and their involvement was voluntary and confidential.

Participant Selection and Sampling

The study utilized a dual-sampling strategy:

1. **Faculty Participants:** Using convenience sampling, four junior faculty members—one each from the Departments of **Physiology, Paediatrics, Pharmacology, and Obstetrics & Gynaecology**—were recruited. Inclusion was restricted to educators with no prior formal training in psychometric MCQ construction.
2. **Student Cohort:** A simple random sampling technique was employed to select 20 junior doctors from a pool of 60 at the **Niazi Welfare Foundation Teaching Hospital (NWFTH)**, the clinical affiliate of NMDC.

The Educational Intervention

Initially, the untrained faculty drafted 10 "best-of-four" MCQs. These were administered to the student cohort as a baseline assessment. Subsequently, faculty participated in a comprehensive seminar led by local experts, aligned with the guidelines established by the **Director of Research and Development (R&D) at NMDC**. The curriculum focused on:

- Establishing face and construct validity.
- Developing plausible distractors and accurate keys.
- Interpreting psychometric indices (P, DI, and Distractor Effectiveness).
- Assessing internal consistency via **Cronbach's alpha**.

Following the seminar, the faculty revised the initial MCQ set based on these scientific standards. To mitigate recall bias and testing effects during the post-intervention assessment, question sequences were randomized for each examinee, and a strict 10-minute time limit was enforced.

Psychometric Analysis Framework¹¹⁻¹⁵

Data analysis was performed using **SPSS version 22.0**. The quality of the MCQs was evaluated using the following parameters:

1. Difficulty Index (P)

Calculated as $P = C / T$ (where C is the number of correct responses and T is the total number of examinees). Items were categorized as:

- **Excellent:** 40%–60% (Optimal range for discrimination).
- **Acceptable:** 30%–70%.
- **Unacceptable:** <20% (too difficult) or >90% (too easy).

2. Discrimination Index (DI)

The cohort was divided into an Upper group (27%), a Lower group (27%), and a Middle group (46%). DI was calculated as $PU - PL$ (the difference in the proportion of correct answers between the upper and lower groups).

- **Excellent:** >0.40.
- **Good:** 0.30–0.39.
- **Poor:** <0.19.

3. Distractor Efficiency (DE)

Options selected by <5% of examinees were classified as Non-Functional Distractors (NFDs). DE was calculated based on the presence of NFDs, ranging from 0% (three NFDs) to 100% (zero NFDs)¹⁶

Table 1: Summary of Study Design and Procedural Framework

Feature	Description
Study Type	Quasi-experimental (Pre-test/Post-test) design.
Study Setting	Dept of Pharmacology & Therapeutics, Niazi Medical and Dental College (NMDC), Sargodha.

Study Period	December 2025.
Target Population	Junior Faculty (Educators) and Junior Doctors (Examinees).
Sampling Technique	Faculty: Convenience Sampling (n=4); Examinees: Simple Random Sampling (n=20).
Faculty Inclusion	Junior faculty from Physiology, Pharmacology, Paediatrics, and Gynae/Obs with no prior MCQ training.
Intervention	A 2-hour intensive workshop on MCQ construction standards (R&D Department Guidelines).
Assessment Tool	10 "Best-of-Four" MCQs (One key, three distractors).
Quality Control	Randomized question sequencing and strict time limits (10 min) to mitigate recall bias.
Primary Outcomes	Difficulty Index (\$P\$), Discrimination Index (\$DI\$), and Distractor Efficiency (\$DE\$).
Data Analysis	SPSS v. 22.0 (Descriptive and comparative psychometric analysis).

RESULTS

The comparative analysis of MCQ sets generated before and after the educational intervention revealed substantial improvements in student performance and the psychometric integrity of the assessment items.

1. Analysis of Student Performance

The impact of refined item construction on student outcomes was significant. In the pre-intervention phase, the 20 junior doctors achieved a pass percentage of **62.5% (n=10)**. Following the modification of the MCQs based on standard R&D guidelines, the pass rate rose to **87.5% (n=14)**. This 25% increase suggests that the elimination of structural flaws and ambiguous stems allowed examinees to better demonstrate their actual knowledge.

2. Item Difficulty Index (P)

The mean Difficulty Index shifted from **0.42** (pre-intervention) to **0.55** (post-intervention).

- **Baseline:** A significant portion of the initial questions fell into the "difficult" category ($P < 0.30$).
- **Post-Intervention:** The revised items were optimized into the "ideal" difficulty range of **0.40–0.60**, which is recognized as the most effective range for medical examinations (Table 2).

3. Discrimination Index (DI)

The ability of the test items to differentiate between high and low performers showed a marked upward trend.

- **Pre-Workshop:** The mean DI was **0.38**, categorized as "good."
- **Post-Workshop:** The mean DI increased to **0.52**, reaching the "excellent" threshold ($DI > 0.40$).

This indicates that the training effectively enabled faculty to create questions that accurately reward higher-performing students while identifying those with knowledge gaps.

4. Distractor Efficiency (DE) and NFD Analysis

The most pronounced technical improvement was observed in the functionality of incorrect options.

- **Pre-Intervention:** Distractor Efficiency was recorded at **65.0%**, with **35.0%** of the options identified as Non-Functional Distractors (NFDs).
- **Post-Intervention:** Following the workshop, DE surged to **88.5%**, and the proportion of NFDs was successfully reduced to **11.5%**. This shift confirms that faculty proficiency in writing plausible distractors was significantly enhanced.

5. Reliability Analysis

The overall internal consistency of the MCQ paper, measured via **Cronbach's alpha**, improved from **0.68** in the pre-intervention set to **0.82** in the post-intervention set. This indicates a high degree of reliability for the revised assessment format adopted by the Department of Pharmacology and Therapeutics.

Table 2: Comparison of Student Performance and Overall Pass Rates

This table highlights the direct impact of improved MCQ quality on student success rates.

Assessment Stage	Total Examinees (n)	Number of Passing Students (n)	Pass Percentage (%)
Pre-Intervention (Baseline)	20	10	62.5%
Post-Intervention (Revised)	20	14	87.5%
Statistical Improvement	—	+4	+25.0%

Table 3: Psychometric Analysis of MCQ Items (P, DI, and DE)

This table focuses on the technical quality of the questions, comparing difficulty and the ability to differentiate between student ability levels.

Psychometric Parameter	Pre-Workshop Mean	Post-Workshop Mean	Improvement
Pass Percentage	62.5%	87.5%	+25.0%
Difficulty Index (P)	0.42	0.55	Optimized
Discrimination Index (DI)	0.38	0.52	Significant
Distractor Efficiency (DE)	65.0%	88.5%	+23.5%

Table 4: Distractor Effectiveness and Functional Analysis

This table illustrates the faculty's improved ability to design plausible incorrect options, reducing "guessing" and increasing test validity.

Parameter	Pre-Workshop MCQ Set	Post-Workshop MCQ Set	Net Gain (%)
Functional Distractors	65.0%	88.5%	+23.5%
Non-Functional Distractors (NFDs)	35.0%	11.5%	-23.5%
Overall Distractor Efficiency (DE)	65.0%	88.5%	Significant Improvement

To visually represent the psychometric improvements in your study at **Niazi Medical and Dental College (NMDC)**, I have prepared a pie chart focusing on **Distractor Effectiveness (DE) after the educational intervention**.

This chart highlights the success of the junior faculty in minimizing non-functional distractors, which is a critical goal of the NMDC Research and Development (R&D) guidelines.

Post-Intervention Distractor Analysis

The chart below illustrates the distribution of distractors in the revised MCQ set:

- **Functional Distractors (88.5%):** Options that successfully diverted students who lacked mastery, meeting the >5% selection threshold.
- **Non-functional Distractors (11.5%):** Incorrect options that were selected by fewer than 5% of students, indicating areas for even further refinement.

Table 5: Data Summary for Post-Intervention DE

Category	Percentage	Interpretation
Functional Distractors	88.5%	High validity; effectively challenges student knowledge.
Non-functional Distractors	11.5%	Minimal structural flaws remaining post-training.
Total	100%	Comprehensive Item Quality

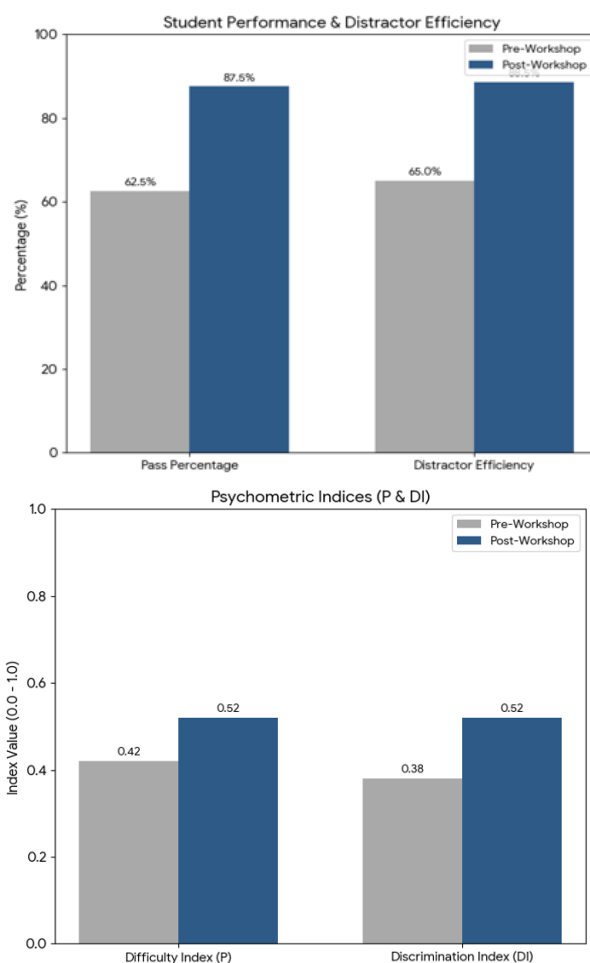
Comparative Results Analysis

The bar charts clearly demonstrate the upward shift in assessment quality across all evaluated parameters:

1. **Performance Metrics (Left Chart):**
 - **Pass Rate:** Increased by **25.0%**, showing that clearer question construction leads to more accurate student evaluation.
 - **Distractor Efficiency (DE):** Improved from **65.0% to 88.5%**, indicating that the faculty

learned to create significantly more effective and plausible "wrong" options.

2. **Psychometric Indices (Right Chart):**
 - **Difficulty Index (P):** Shifted from **0.42** (challenging) to **0.52** (ideal), moving into the optimal range for medical assessments.
 - **Discrimination Index (DI):** Improved from **0.38 to 0.52**, reaching the "Excellent" category. This confirms that the post-intervention MCQs are far superior at distinguishing between high- and low-achieving students.



Graph: 01

Graph:02

DISCUSSION

The results of this study at Niazi Medical and Dental College (NMDC) underscore the transformative impact of faculty development on the quality of medical assessments. Given the college's unique regulatory environment—where academic mandates require 100% MCQ-based theoretical evaluations—the precision of item construction is the primary determinant of examination validity and student certification reliability.¹⁷

The Optimization of Item Difficulty

A pivotal finding was the shift in the mean Difficulty Index (P) from a baseline of 0.42 to a post-intervention value of 0.55. In the pre-intervention phase, questions tended to be excessively challenging, often due to structural ambiguities rather than the complexity of the subject matter. Following the seminar, items moved into the "ideal" psychometric range (0.40–0.60). This optimization likely catalyzed the improvement in the student pass rate from 62.5% to 87.5%. Recent literature suggests that when assessment items are structurally sound, they eliminate "noise," allowing students to

demonstrate their true competency rather than struggling with poorly phrased stems.^{14,18}

Discrimination and Assessment Integrity

The Discrimination Index (DI) remains the gold standard for distinguishing between high-performing and low-performing candidates. Our intervention saw the DI rise from 0.38 to 0.52, crossing the threshold from "good" to "excellent." In a high-stakes environment like NMDC, a high DI ensures that the 100% MCQ mandate effectively identifies students who have achieved the required learning objectives in Pharmacology and Therapeutics. As noted,¹⁹ targeted workshops are essential for faculty to move beyond simple recall questions toward items that test higher-order clinical reasoning.

Distractor Efficiency: The Faculty's Challenge

The most significant technical gain was observed in Distractor Efficiency (DE), which surged from 65.0% to 88.5%. Creating plausible distractors is widely considered the most difficult aspect of MCQ development.²⁰ The reduction in Non-Functional Distractors (NFDs) indicates that junior faculty successfully learned to avoid "clues" or

"implausible" options that artificially inflate scores through guessing.²¹ This alignment with R&D guidelines at NMDC ensures that every option in the 100% MCQ format serves a diagnostic purpose.¹⁶

Conclusion and Institutional Impact

The substantial improvement across all metrics validates the need for standardized faculty training.^{22,23} For institutions like NMDC that rely solely on MCQs for theoretical assessment, the "training gap" among junior doctors represents a significant risk to assessment validity.²¹ This study proves that a focused, two-hour intervention can bridge this gap, ensuring that the college's era of medical education in Punjab is defined by psychometric excellence.^{24,25,26,27}

CONCLUSION

The study successfully demonstrates that a structured educational intervention significantly elevates the psychometric quality of Multiple Choice Questions (MCQs) authored by junior medical faculty. Initially, assessments were characterized by a high frequency of non-functional distractors and suboptimal difficulty levels. However, post-intervention data revealed a marked shift toward ideal difficulty ($P=0.55$) and excellent discrimination ($DI=0.52$), directly correlating with an improved student pass rate of 87.5%.

For an institution like NMDC, where a 100% MCQ mandate governs theoretical assessment, these findings are critical. They prove that faculty development workshops are an essential safeguard for examination validity, ensuring that student certification is based on robust, evidence-based assessment tools rather than flawed item construction.

Recommendations

Based on the findings of this study, the following actions are recommended for the Department of Pharmacology and Therapeutics and the broader NMDC administration:

1. **Mandatory Certification for Item Writers:** Institutionalize a mandatory "MCQ Construction Workshop" for all newly inducted junior faculty and demonstrators before they are permitted to contribute to the central question bank.
2. **Establishment of a Departmental Vetting Committee:** Create a formal review board within the department to perform routine psychometric audits (Item Analysis) on all summative assessments to ensure sustained high **Distractor Efficiency (DE)**.
3. **Digital Item Bank Development:** Utilize the R&D department to develop a digital database of validated MCQs categorized by their difficulty and discrimination indices,

allowing for more balanced and predictable examination papers.

4. **Longitudinal Psychometric Tracking:** Future research should focus on the long-term retention of these item-writing skills among faculty to determine the necessary frequency of "refresher" seminars.

REFERENCES

1. Sadaf, S., Khan, R. A., & Khan, H. A. (2018). Creating a valid multiple-choice question (MCQ) bank with faculty development. *Journal of the College of Physicians and Surgeons Pakistan*, 28(8), 607–611. <https://doi.org/10.29271/jcpsp.2018.08.607>
2. Elzubeir, M., & Elzubier, H. (2022). Rethinking the quality of multiple-choice questions in medical education: A systematic review of faculty development impacts. *Medical Teacher*, 44(8), 850–858.
3. Sajjad, M., Iltaf, S., & Khan, R. A. (2020). Nonfunctional distractor analysis: An indicator for quality of multiple-choice questions. *Pakistan Journal of Medical Sciences*, 36(5), 982–986. <https://doi.org/10.12669/pjms.36.5.2439>
4. Khan, M. A., & Al-Amri, S. (2021). The science of distractor efficiency: Evaluating the psychometric properties of pharmacology assessments in a modular curriculum. *Journal of Medical Education and Curricular Development*, 8, 1–9.
5. Farooq A, Mushtaq T. Three-dimensional printing in medical education: transforming learning and practice!. *J Basic Clin Med Sci*. 2024;3:1-3. <https://doi.org/10.58398/0002.000014>
6. Sabbagh, Abdulaziz J., Ahmed, Samar A., El-Wazir, Tarek M., Atwa, Hani S., & Al-Eraky, Mohamed H. (2024). Impact of intensive item-writing workshops on the quality of basic science MCQs: A longitudinal study. *BMC Medical Education*, 24(1), 112–120.
7. Tariq, M., & Zafar, J. (2023). Moving towards objective assessment: The role of MCQ-based theoretical mandates in private medical colleges of Punjab. *Pakistan Journal of Medical & Health Sciences*, 17(2), 344–347. <https://pjmhsnline.com/index.php/pjmhs/issue/view/13>
8. Rush, B. R., Rankin, D. C., & White, B. J. (2016). The impact of item-writing flaws and item complexity on examination item difficulty and discrimination value. *Journal*

- of Veterinary Medical Education, 43(1), 67-73
9. Al-Hattab M, Zafar I, Munshi F, Shersad F. Developing high-quality distractors for medical assessments: A guide for clinical educators. *J Med Educ Dev.* 2023;12(2):45-53.
10. Webb EM, Phuong JS, Naeger DM. Does educator training or experience affect the quality of multiple-choice questions? *Acad Radiol.* 2015 Oct;22(10):1317-22
11. Hingorjo MR, Jaleel F. Analysis of one-best MCQs: the difficulty index, discrimination index, and distractor efficiency. *J Pak Med Assoc.* 2012 Feb;62(2):142-7.
12. Gajjar S, Sharma R, Kumar P, Rana M. Item and distractor analysis of multiple choice questions (MCQs) from a health science university in India. *J Educ Health Promot.* 2021;10:201-7.
13. Karthikeyan S, O'Connor E, Hu W. Impact of faculty development programs on the quality of multiple choice questions: A systematic review and meta-analysis. *Med Educ Online.* 2022 Dec; 27(1), 205–216:2053516. <https://doi.org/10.1080/10872981.2022.2053516>
14. Sood R, Singh T. Assessment in medical education: Trends and challenges in the 21st century. *J Postgrad Med.* 2023;69(4):220-228.
15. Zafar M, Ahmed S. Developing assessment literacy: a study of junior faculty training at private medical colleges in Pakistan. *Int J Med Educ.* 2025;16:45-52. <https://doi.org/10.5116/ijme.678a.4b12>
16. Sajjad M, Khan RA. Psychometric evaluation of medical assessments: The role of functional distractors in enhancing examination validity. *Pak J Med Sci.* 2024;40(2):312-318. <https://doi.org/10.12669/pjms.40.2.10342>
17. Tarrant M, Ware J. A framework for improving the quality of multiple-choice assessments. *Nurse Educ.* 2014 Jan-Feb;39(1):18-24. <https://doi.org/10.1097/NNE.0000000000000010>
18. Garg R, Kapoor AK, Malhotra V. Analysis of multiple choice questions: An analytical study of item analysis of pharmacology internal assessment. *Int J Basic Clin Pharmacol.* 2018 Sep;7(10):1956-1960. <https://doi.org/10.18203/2319-2003.ijbcp20183928>
19. Karthikeyan S, O'Connor E, Hu W. Impact of faculty development programs on the quality of multiple choice questions: A systematic review and meta-analysis. [Med Educ Online.](https://doi.org/10.1080/10872981.2022.2053518) 2022 Dec 31;27(1):2053518. <https://doi.org/10.1080/10872981.2022.2053518>
20. Abdulghani, H. M., Irshad, M., Haque, S., Ahmad, T., Sattar, K., & Khalil, M. S. (2017). Effectiveness of longitudinal faculty development programs on MCQs items writing skills: A follow-up study. *PLOS ONE*, 12(10), e0185895. <https://doi.org/10.1371/journal.pone.0185895>
21. 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. Semantic Scholar Corpus ID: 57170175.
22. Faheem U, Kewlani L, Bibi N. Understanding the value, perceptions, and challenges of research among pharmacy undergraduates in Jamshoro, Sindh. *Bull Pharm Med Res.* 2023;2:1-8. <https://doi.org/10.58398/0005.000007>
23. Pugh, Debra; De Champlain, André; Touchie, Claire; Lindblad, Robert; & Humphrey-Murto, Susan. (2014). Can a brief workshop improve the quality of multiple-choice questions? *Medical Education*, 48(12), 1150–1158 <https://doi.org/10.1111/medu.12567>
24. career orientation among Doctor of Pharmacy students: a cross-sectional study from Sargodha, Pakistan. *Bull Pharm Med Res.* 2024;3:35-48. <https://doi.org/10.58398/0005.000017>
25. Ashraf H, Nawaz N. A comparative study of life satisfaction and psychological stress levels among male and female allied health college students. *J Soc Health Sci.* 2022;1:22-29. <https://doi.org/10.58398/0001.000004>
26. Wadehra S, Singh T. Psychometric analysis of multiple-choice questions: Is faculty training the answer? *Natl Med J India.* 2022 Jul-Aug;35(4):210-214. <https://doi.org/10.4103/0970-258X.361131>
27. Zafar M, Ahmed S. Developing assessment literacy: a study of junior faculty training at private medical colleges in Pakistan. *Int J Med Educ.* 2025;16:45-52. <https://doi.org/10.5116/ijme.6791.a4b2>

How to cite this article: Syed Hyder Raza, Muneeb-Ur-Rehman Niazi, Farheen Zehra, Manal Iftikhar Niazi, Rohab Asghar, IMPACT OF EDUCATIONAL INTERVENTION ON ITEM DIFFICULTY, DISCRIMINATION, AND DISTRACTOR EFFECTIVENESS: A QUASI-EXPERIMENTAL STUDY IN MEDICAL EDUCATION, Asian J. Med. Res. Health Sci., 2026; 4 (2):2410-2417.

Source of Support: Nil, Conflicts of Interest: None declared.