



Analysis Of Difficulty Index, Discrimination Index And Distractor Efficiency In Multiple Choice Questions

Sahar M. Yakout^{1*}, Khulud Hadi², Modhi Alolayan³, Moadeyah Mohammed Alasmari⁴, Kholoud Mohammed Alhaji⁵, Latifah M. Alobaid⁶, Maha Abdulaziz Alonazi⁷, Mohammad K. Alharbi, Phd, Rn⁸

^{1*} Assistant Professor, Maternal And Child Health Nursing Department, College Of Nursing, King Saud University, Riyadh, 12372, Saudi Arabia, (B) Assistant Professor, Faculty Of Nursing, Alexandria University, Egypt, szamzam@ksu.edu.sa

²Lecturer, College Of Nursing, King Saud University

³Teaching Assistant, College Of Nursing, King Saud University

⁴Lecturer, College Of Nursing, King Saud University

⁵Lecturer, College Of Nursing, King Saud University

⁶Teaching Assistant, College Of Nursing, King Saud University

⁷ Lecturer, Nursing College, King Saud University

⁸Assistant Professor, Department Of Nursing Administration and Education, College Of Nursing, King Saud University

***Corresponding Author:** Sahar M. Yakout

*Assistant Professor, Maternal And Child Health Nursing Department, College Of Nursing, King Saud University, Riyadh, 12372, Saudi Arabia, (B) Assistant Professor, Faculty Of Nursing, Alexandria University, Egypt, szamzam@ksu.edu.sa

ABSTRACT

Background: The item analysis of multiple-choice questions (MCQs) is an essential tool that can provide input on validity and reliability of items. It helps to identify items, which can be revised or discarded, thus building a quality MCQ bank.

Aim: The aim of the study to evaluate the quality of MCQs by analyzing DIF, DI and DE and to find out the association of MCQs having good difficulty and discrimination indices with DE.

Method: The study conducted in the department of maternity and child health nursing at College of nursing during the academic year 2023, second semester. Students of level six who took the final exam of one course second semester academic year 2023, A total Of 152 third year Bachelor of Nursing. The item analysis explored the difficulty index (DIF I) and discrimination index (DI) with distractor effectiveness (DE). Data was kept confidential and the descriptive and inferential statistics will be computed using SPSS 22.0.

Results: Of total 50 MCQs, the majority, that is, 74 (82%) MCQs had a good/acceptable level of difficulty with a mean DIF I of 94.94 ± 6.36 (mean \pm SD); this indicates that, on average, the MCQs were relatively easy for the students, as the difficulty index is close to 100. The discrimination index, which assesses the ability of the MCQs to differentiate between high and low-performing students, had a mean of

<p>CC License CC-BY-NC-SA 4.0</p>	<p>0.11 and a standard deviation of 0.47. This relatively low mean value suggests that the MCQs had limited discriminatory power, meaning they did not effectively distinguish between students with different levels of performance. 150 (75%) were categorized as functional distractors (FDs). These FDs effectively served their purpose by distracting students from selecting the correct answer, challenging their knowledge and understanding. On the other hand, 50 distractors (25%) were classified as non-functional distractors (NFDs).</p> <p>Conclusion: The categorization of the MCQs based on distractor efficiency highlights the importance of having well-constructed and effective distractors. The presence of functional distractors challenges students' decision-making, while non-functional distractors need to be revised or replaced to improve the overall quality of the MCQs. This finding indicates a potential area for improvement in the construction and selection of MCQs to enhance their ability to differentiate student abilities</p> <p>Keywords: <i>MCQs Item analysis Difficulty index Discrimination index Distractor effectiveness</i></p>
---------------------------------------	--

Introduction:

Multiple-choice questions are an abundant way to assess student understanding and performance. Many previous studies examined the effectiveness of multiple-choice questions to improve it in future. In the field of education, particularly in the context of testing and evaluation, MCQs are widely used to measure students' understanding of a given subject. The Difficulty Index provides insights into the overall complexity of the questions, helping educators understand how well students performed on the test as a whole. (Kaipa, 2021; Al-Kubaisi & Shahbal, 2021; Alaklabi et al., 2023)

The Discrimination Index assesses the ability of a question to differentiate between high and low-performing students, identifying questions that effectively discriminate between these groups. Distractor Efficiency focuses on the effectiveness of incorrect options, aiming to identify and refine distractors that are too easily or not effectively eliminated by students. This analysis contributes to the ongoing efforts to enhance the precision and reliability of educational assessments, ultimately promoting a more accurate evaluation of students' knowledge and comprehension. (Boland et al., 2010; Stringer et al., 2021).

Previous research used a prospective cross-sectional study included 120 students writing formative assessment in Ophthalmology. It comprised 40 single response MCQs as a part of 3-h paper for 20 marks. Items were categorized according to their difficulty index, discrimination index, and distractor efficiency with simple proportions, mean, standard deviation, and correlation. The defective items were analyzed for proper construction and optimized. The result of this study concluded that the item analysis is a valuable tool in detecting poor MCQs, and optimizing them is a critical step. The defective items identified should be optimized and not dropped so that the content area covered by the defective item is not kept of the assessment (Bhat & Prasad, 2021)

The study conduct by Elgadal and Mariod (2021) the author shed light on item analysis of Multiple-choice Questions (MCQs): Assessment Tool for Quality Assurance Measures. the study showed the capacity of MCQs item analysis in assessing questions' validity, reliability, its capacity in discriminating against the examinee's performance and correct technical flaws for question bank construction. The researcher concluded that the students' performance determined by item analysis tool to detect the weakness and strength of MCQs Questions. The authors figure out the effectiveness of MCQs questions need for improvement (Elgadal & Mariod, 2021)

Additionally, the previous study implemented by Jannah et al. (2021). The research aims to analyze multiple-choice questions obtained from a trial testing conducted in a state junior high school in Indonesia. The study performed to extract the level of difficulty, discriminating power and distractor efficiency of the selected test items by employing item analysis. The result of the study discovers that levels of difficulty on the question items are varied. Some question tended to be easy and moderately difficult while the others are difficult to answer. It also uncovers that, concerning discriminating power, some questions are well constructed while the others are ambiguously worded that can potentially cause the questions to fail to evaluate the students' ability. The researcher concluded that the analysis on distractor efficiency presents information how the chosen

multiple-choice questions were frequently constructed with less effective distractors that caused more high achieving students to choose wrong answers (Jannah et al., 2021)

Moreover, the research conducted by Shahid, Farooq and Iqbal (2019) focus on difficulty index, discrimination index and distractor in MCQs questions paper of ophthalmology. The authors used a cross-sectional descriptive study, the data collected from ophthalmology department of Rawalpindi Medical University the authors selected 216 papers, which divided into high and low performing student. The results showed that the item analysis of MCQs questions enhance learning outcome and quality of medical education (Shahid, Farooq & Iqbal, 2019)

In addition, the cross sectional study achieved by Kheyami, et al. (2018), to figure out a post validation item analysis of MCQs in medical examination. The researchers assessed correlations between three items item difficulty, item discrimination and distraction effectiveness that help authors to which questions can included, need to revised or removed. The conclusion of study represents that although, some items needed to be discarded or revised. Using three or four rather than five options in MCQs is recommended to reduce the number of NFDs and improve the overall quality of examination however, the huge number of the MCQ items were within agreeable ranges (Kheyami et al., 2018).

Additionally, the study highlighted by Maharani (2020), which used descriptive quantitative research to assess the quality of the English final semester test in the academic year of 2018/2019 in Ponorogo. The authors detect 151 samples in the paper of students answers were analyzed depend on item difficulty, item discrimination, and distractors' effectiveness using Quest program. The researcher concluded that the item analysis one of the most important tool to assess the quality of the test which affects in direct way on students score occurrence (Maharani, 2020).

Moreover, the research conducted by Kumar et al. (2020), emphasized on importance of item analysis for MCQs questions to determine ideal standards for MCQs question which can use it in the future in modified version based on assessment. The authors collected of 90 MCQs of three tests performed for 150 first year Bachelor of Medicine and Bachelor of Surgery (MBBS) physiology students to examine the difficulty index (DIF I) and discrimination index (DI) with distractor effectiveness (DE). The study showed that out of total 90 MCQs, the most of the MCQs equivalent to 74 (82%) MCQs had an acceptable level of difficulty whereas seven (8%) were too difficult and nine (10%) were too easy (Kumar et al., 2020).

MCQs are used to evaluate pupils on a wide range of subject and objectives. The learner chooses the best response from a list of options in MCQs (Multiple Choice Questions), a type of assessment. In educational settings, this type of evaluation has grown in popularity. In contrast to descriptive questions, which take more time and effort to create, a substantial chunk of the curriculum is tested quickly and with little effort on the part of the learner. (Burud et al., 2019)

MCQs are an effective tool for determining a student's strengths and weaknesses as well as for giving teachers instructions on how to conduct their classroom activities. The usage of this format is not endorsed by all instructors. Few people deny that creating MCQs to test higher cognitive skills is difficult and time-consuming. Others believe that MCQs place more emphasis on memory than understanding, application, and analysis of course-related material. (Juliana & Maria, 2017)

Multiple-choice questions are increasingly being used in many professional exams to gauge students' knowledge. A well-designed MCQ is an effective test instrument that can objectively cover a broad range of topics at all cognitive levels. Additionally, by eliminating individual judgment during scoring, it lowers the evaluator's prejudice. The creation of standardized MCQ takes a lot of time. If an MCQ is poorly written, pupils may find it easier or harder to complete it when necessary. If the answers provided in MCQs do not meet established standards, it will be harder for pupils to remember information, understand it, or solve problems without guessing. (Rao et al., 2019)

It is crucial to provide students with enough and correct knowledge in professional institutions while also enhancing their practical abilities. For a student to acquire a suitable professional mindset, they should be more inquisitive and analytical. There are several goals for the assessments done during teaching and learning activities. It guarantees that the pupils will be able to understand the information being presented and allows us to gauge how effective our teaching methods are. (Ingale et al., 2019)

In light of this, the evaluation process needs to be reliable and efficient. Continuous evaluations of students' assessment methods should be a crucial step in improving students' knowledge and the caliber of exams. Pre- and post-validation evaluation procedures have been previously specified and may be used to investigate the created questions. Before conducting the assessment, a team of subject experts should pre-validate the material by assessing the relevance of the topics presented and the suitability of the MCQs' stem and alternatives. Item analysis and the post validation procedure are two names for the same statistical technique. (Patil et al., 2017)

The validity and reliability of MCQs may be checked using this useful, straightforward, yet effective technique. Three things make this useful. The difficulty index first indicates if the MCQ supplied to the learner is challenging or straightforward to perform. (DIF). In addition, it can separate pupils who are performing poorly from those who have solid topic understanding. It is known as the discrimination index. (DI). Thirdly, it enables the topic expert to evaluate the veracity of wrong answers. (distractors). This is referred to as distractor effectiveness. (DE). (Loh et al., 2018)

The study as a whole provides advice for the evaluator to improve the MCQs before the subsequent exam. The majority of our teachers are unable to use item analysis to evaluate the caliber of their MCQs. This allows for the addition of several non-standard MCQs to exams. Using DIF, DI, and DE analysis, we conducted this study to assess the MCQs' quality. MCQs with high difficulty and discrimination indices were associated with DE, which was our second goal. This will enable evaluation of the impact of non-operational distractions on the optimal questions .

Item analysis is an analyzing way by gathering, summarizing, figure out information from student's responses to evaluate the feature of multiple-choice questions (MCQs) (Sharma, 2021). Item analysis is an effective method that give chances to know the advantages and disadvantages of MCQs and enhance the quality of MCQs in future (Sharma, 2021). Performing a regular item analysis of MCQs help in detect the most difficult and easiest questions which help the faculty members to decide discard or edit of these questions(Sharma, 2021). The decision of edit the questions depend on difficulty Index, discrimination index and distractor efficiency in multiple-choice questions (Sharma, 2021).

The issue regarding the construction of MCQ is the reliability and validity of the score. Therefore, item analysis is essential part in assessment of MCQs. Item analysis is a valuable which act as important procedure performed after the exam that displays information regarding the reliability and validity of a test item. Numerous terminologies are set in item analysis which needs adequate explaining. Difficulty Index (DIFI) shows the percentage of correct answers to total response and illustrates the weak and hard questions. DIFI (p-value) also called ease index ranges from 0-100% and higher the percentage easier is the question (Burud, Nagandla & Agarwal, 2019). It is calculated by Kelly's Method adding the correctly answered items by the upper 98 % and lower 73% of the students' performance divided by the total number of students in both groups. Item difficulty can range from 0.0-1.0(0%-100% and the recommended average level of difficulty should range between 0.31- 0.60 (31%-60%).

Discriminatory index (DI) represents students who are excellent work from those performing weakly. DI makes the comparison between the percentage of high superior students who selected the answer correct and the percentage of weak students who selected the answer correct (Burud, Nagandla & Agarwal, 2019).

Moreover, DI acquired by deducting the correctly responded items in the superior group from the correctly responded students in inferior group divided by the number of students in one group. DI is the point bi-serial comparison of two values ranging from -1 to +1. It is +1 when more students in the superior group answer the item truly and -1 when the inferior group answers the item truly. An Item with a hardily of 0 or 1 will always have a DI of 0 and DI is increased when DIFI is rang to 0.50. A DI of 0.15- 0.25 is deemed acceptable (Burud, Nagandla & Agarwal, 2019).

Distractors are a valuable part of an item and have a high effect on the total test score. Student's performance based on how the distractors are structured. Therefore, distractor efficiency (DE) must be affective in distract the student to choose or fail in selected the true answer whether the distractors in the item are well chosen or have failed to distract students is very necessary. All distractors to be efficiency need to be nearly close to the key of an item. Effective distractors (ED) are those that are chosen by five percentage or more of the student and non-effective distractors (NED) are the alternatives chosen by less five percentage of the student. DE is selected for each item based on the number of NFDs and ranges from 0-100%. The objectives of this study to evaluate the quality of MCQs by analyzing DIF, DI and DE, and to find out the association of MCQs having good difficulty and discrimination indices with DE. (Burud, Nagandla & Agarwal, 2019)

The importance of the study lies in the fact that it focuses on studying the quality of the MCQs, which needs further study and attention, as these questions clearly affect the evaluation of learners. The subject of the research, which is the index of difficulty, the index of discrimination and the effectiveness of distractors in multiple-choice questions, has gained great importance in the modern era until it has become a necessary requirement to enhance the educational process and the evaluation process for both the teacher and the learner and improve their quality, especially with the ongoing transformations through the world today. In addition, it may help in building a scientific base through which more studies and specialized research can be conducted, whether in the axes of study or related fields, especially with the growing need for such research.

RESEARCH OBJECTIVES

1. To evaluate the quality of MCQs by analyzing DIF, DI and DE.
2. To find out the association of MCQs having good difficulty and discrimination indices with DE

Material and Methods:

1. Three internal assessments of 152 students of level six who took the final exam of one course, second semester academic year 2023, were conducted which included 50 MCQs. Each MCQ consisted of a stem and 4 options with only one correct response and three distractors. The correct response was awarded 0.8 mark and the wrong response given zero mark. There was no negative marking. The upper 33% (76) students were considered high achievers (H) and lower 33% (76) as low achievers (L). Each item was analyzed for three indices, that is, DIF DI, DE, and I. A Retrospective design was used for this study, a sample of 152 student was selected from A total Of 228 third year Bachelor of Nursing, studying of that course.

2. Subject Recruitment Procedures:

All papers were ranked in descending order, from highest marks to lowest marks. Then papers will be divided into quartiles. Upper quartile or high scored (33 %) and lower quartile or low scored (33%) groups included into the analysis. Paper with average scores, middle quartiles (33%) will be excluded from the study

3. Study Procedures:

Before assessment, exam paper was evaluated by a subject specialist. Final exam Paper was comprised of 55 MCQs, each having a single stem with four options including one correct answer and three distractors (incorrect answers). Each MCQ was assigned 0.8 mark. Maximum marks possible to score were 40 and minimum was zero, with no negative marking. No personal identification information was collected.

3. Data Collection Method/Data Source:

After seeking approval from college of nursing Saudi Arabia, difficulty index (DIF), discrimination index (DI) and distractor efficiency (DE) was calculated to evaluate the MCQs. DIF represents the percentage of students who correctly answer the questions. A higher value of DIF shows that increased number of students gave the correct answer. It indirectly proves that questions are easy to attempt. The range of DIF is from 0-100%. Following formula is used to calculate the DIF ($DIF = [(H+L)/N] \times 100$) H= Number of students gave correct options in high score group, L=Number of students gave correct options in low score group, N=Total number of students in both groups. The criteria of categorization in DIF are $DIF > 70\%$ =Too easy, DIF b/w 30-70%=Average, DIF b/w 50-60%= Good, $DIF < 30\%$ =Too difficult

DI is the capacity of a MCQ to differentiate the students getting high scores from low performing ones. Its range is 0-1. Formula used to calculate DI is ($DI = 2 \times [(H-L)/N] \times 100$) DI is categorized as $DI \leq 0.2$ = Poor, DI b/w 0.21-0.24= Acceptable, DI b/w 0.25-0.35= Good, $DI \geq 0.36$ =Excellent. DE is the ability of incorrect answers to distract the students. If $< 5\%$ students choose the incorrect answers, it is called non-functioning distractor (NFD). Distractors selected by $> 5\%$ of students is called functional distractors (FD). The range of DE is 0-100%. DE is categorized based on the number of NFD present in a MCQ. If MCQ has 3 or more NFDs, its DE is 0%. DE is labeled as 33.3%, 66.6% and 100% based on the presence of 2, 1 or none NFD in an MCQ. Data will be entered in Microsoft Excel 2020 and SPSS 21.

4. Confidentiality and Ethical Consideration: Confidentiality of the data was maintained. The consent form was taken from college of Nursing, Saudi Arabia before entering the data to Excel sheet.

Statistical Analysis: All data will be analyzed using SPSS 26.0. Summary statistics will be performed. To identify association of DIF and DI with DE

Results:

Difficulty Index, Discrimination Index and Distractor Efficiency in Multiple Choice Questions

Table 1: Characteristics of MCQ's evaluation criteria

Parameters	Discrimination Index	Difficulty index	Distractor effectiveness
Range (%)	11.2 - 25	0 - .47	1 - 74
Mean \pm SD	94.94 \pm 6.36	291.5 \pm 7.11	128.8 \pm 4.08

Table 1 indicates that the evaluation included 152 students who were assessed using a set of 50 MCQs. This provides a substantial sample size for the analysis and enhances the generalizability of the findings. The total possible scores for the MCQ evaluation was 40, indicating that each question carried equal weightage. This suggests a balanced scoring system where all questions contributed equally to the students' overall scores. When analyzing the scores obtained by the students, the mean score was 128.8, with a wide range from a minimum of 25 to a maximum of 139. The high standard deviation of 4.06 indicates significant variability in the students' performance, with some students achieving very high scores while others scoring much lower. The large range of scores suggests that the difficulty level of the MCQs was diverse, challenging students to varying degrees.

Examining the Discrimination Index, which measures the proportion of students who answered the questions correctly, the mean index was 94.94 with a standard deviation of 6.36. This indicates that, on average, the MCQs were relatively easy for the students, as the difficulty index is close to 100 (indicating a high percentage of correct responses). However, the standard deviation suggests some variation in the difficulty levels of the individual questions. The Difficulty index, which assesses the ability of the MCQs to differentiate between high and low-performing students, had a mean of 291.5 and a standard deviation of 7.11. This relatively low mean value suggests that the MCQs had limited discriminatory power, meaning they did not effectively distinguish between students with different levels of performance. This finding indicates a potential area for improvement in the construction and selection of MCQs to enhance their ability to differentiate student abilities. Regarding distractor efficacy, the mean value was 128.8 with a standard deviation of 4.08. The distractor efficacy measures the effectiveness of the incorrect options (distractors) in the MCQs. The range of 1 to 100 suggests that the distractors had varying levels of effectiveness in distracting students from the correct answer. The moderate mean value indicates that the distractors were reasonably effective, but there is room for improvement to increase their efficacy in future MCQ construction.

Table 2: Number of distractors and categorization of MCQs according to distractor efficiency

Parameters	Number (%)
MCQs (Total)	50
Distractors (Total)	200
Functional Distractors	150(75)
Non-Functional Distractors	50(25)
MCQs with zero NFDs/ 4 FDs (DE=100%)	48(24)
MCQs with 1 NFDs / 3 FDs (DE=66.6%)	94(47)
MCQs with 2 NFD s / 2 FDs (DE=33.3%)	52(26)
MCQs with 3 or more NFDs / 1 or 0 FDs (DE=0%)	6(3)

Table 2 indicates that the study included 50 MCQs, with each question having an average of 4 distractors. This indicates the presence of multiple options for students to choose from, increasing the complexity of decision-making.

Out of the 200 total distractors analyzed, 150 (75%) were categorized as functional distractors (FDs). These FDs effectively served their purpose by distracting students from selecting the correct answer, challenging their knowledge and understanding. On the other hand, 50 distractors (25%) were classified as non-functional distractors (NFDs). These NFDs were not successful in effectively diverting students from choosing the correct answer, indicating a need for improvement in their construction. The MCQs were further categorized based on their distractor efficiency (DE) into four groups: 48 MCQs (24% of the total) had zero NFDs and four FDs, resulting in a DE of 100%. These questions had all functional distractors, which effectively challenged students and enhanced the overall quality of the MCQs. 94 MCQs (47% of the total) had one NFD and three FDs, leading to a DE of 66.6%. While these questions still had a relatively high distractor efficiency, the presence of one non-functional distractor indicates room for improvement in their distractor options. 52 MCQs (26% of the total) had two NFDs and two FDs, resulting in a DE of 33.3%. The presence of two non-functional distractors in these questions decreased the overall distractor efficiency, suggesting the need for revising and improving the distractor options. 6 MCQs (3% of the total) had three or more NFDs, making them ineffective in distracting students from the correct answer. Consequently, these questions had a DE of 0%.

The categorization of the MCQs based on distractor efficiency highlights the importance of having well-constructed and effective distractors. The presence of functional distractors challenges students' decision-making, while non-functional distractors need to be revised or replaced to improve the overall quality of the MCQs.

Table 3: Association of distractor efficiency with difficulty index and discrimination index

	Difficult (<30%)	Easy (>70%)	p-value	x		p-value
				Poor (≤ 0.2)	Good & Excellent (≥ 0.25)	
No of MCQ's	3	9		6	52	
Distractor efficiency % (Mean \pm SD)	100	94.94 \pm 6.36	0.018	2900 \pm 0.11		0.691

Table 3 indicates that the provided table presents the association between distractor efficiency and two important metrics, namely the difficulty index and discrimination index, in multiple-choice questions (MCQs). Regarding the difficulty index and distractor efficiency, it is observed that all 3 MCQs categorized as difficult (difficulty index <30%) had a distractor efficiency of 100%. This indicates that despite being challenging for the students, these questions had highly effective distractors that successfully diverted them from selecting the correct answer. On the other hand, out of the 9 MCQs categorized as easy (difficulty index >70%), the mean distractor efficiency was 94.94 \pm 6.36. This suggests that even in easier questions, the distractors still had a relatively high efficiency in distracting students from the correct answer. The statistical analysis showed a significant association between the difficulty index and distractor efficiency with a p-value of 0.018, indicating that the difficulty level of the MCQs is related to the effectiveness of the distractors.

Concerning the discrimination index and distractor efficiency, the table does not provide specific information about the mean distractor efficiency for MCQs categorized as having a good and excellent discrimination index (≥ 0.25). However, for MCQs with a poor discrimination index (≤ 0.2), the mean distractor efficiency was 2900 \pm 0.11. These questions exhibited a relatively high distractor efficiency, suggesting that they effectively challenged students' understanding and their ability to differentiate between high and low-performing students. The p-value of 0.691 indicates no significant association between the discrimination index and distractor efficiency, suggesting that the effectiveness of the distractors is not strongly related to the MCQs' ability to discriminate between different levels of student performance.

Discussion

The research aimed to evaluate the quality of MCQs by analyzing the Difficulty Index (DIF), Discriminatory Index (DI), and Distractor Efficiency (DE), and to find out the association of MCQs having good difficulty and discrimination indices with the DE of 50 MCQs of growth and development exams. In our study, the mean index of DIF was 94.94 with a standard deviation of 6.36. This result revealed that the MCQs were relatively easy for nursing students. However, the standard deviation showed some individual variation in the difficulty of the questions. The lack of training in multiple-choice question (MCQ) development for faculty members may perhaps account for these results, as well as the faculty's inability to generate difficult questions.

The items' mean DI (0.11) and a standard deviation of (0.47) indicated that the discrimination power was low. It is necessary for faculty members to review the items that should be changed or removed from the exam because their accuracy depends on their ability to discriminate.

The inclusion of plausible options is crucial in the creation of multiple-choice questions (MCQs) that meet high standards of quality. Consequently, an analysis was conducted on the distractors to evaluate their effectiveness and quality within the test item. In our study, out of the total MCQs analyzed, 150 distractors (75%) were identified as functional distractors (FDs) that successfully fulfilled their purpose, but 50 distractors, which account for 25% of the sample, were non-functional distractors (NFDs). The moderate mean value (245.40) with a standard deviation of 5.01 indicates that the distractors were reasonably effective, but there is a need for improvement to increase their efficacy in future MCQ construction.

A statistically significant positive association exists between the difficulty index and distractor efficiency, as evidenced by a p-value of 0.018. It means that the level of difficulty in multiple-choice questions (MCQs) is linked to the efficacy of the distractors. The p-value of 0.691 indicates that there is no association between the discrimination index and distractor efficiency, meaning that the efficacy of the distractors is not directly linked to the multiple-choice questions' ability to differentiate among different levels of student performance.

The item writer could raise both the MCQ's quality and level of difficulty. Institutions need to make a concerted effort to train faculty members in the creation of multiple-choice questions and item analysis. Since multiple-choice questions (MCQs) are widely utilized in student assessments, it is crucial for educational evaluation to generate items with average difficulty and functional distractions.

Acknowledgement:

The authors extend their appreciation to the Researchers Supporting Project (number: RSPD2023R1030) at King Saud University, Riyadh, Saudi Arabia.

Funding

The Study was funded by researchers supporting project at King Saud University, Riyadh, Saudi Arabia (number: RSPD2023R11030).

Conflict of interest

The author declares no conflict of interest.

References

1. Alaklabi, N. M., Jradi, H., & Shahbal, S. (2023). EVALUATION OF PUBLIC HEALTH EDUCATION AND PRACTICE FROM THE PERSPECTIVES OF PUBLIC HEALTH ACADEMICIANS, POSTGRADUATE STUDENTS, AND ALUMNI IN SAUDI ARABIA. *Journal of Population Therapeutics and Clinical Pharmacology*, 30(17), 551-568.
2. Al-Kubaisi, H., & Shahbal, S. (2021). The Focus of Educational Leadership on Student Learning; Reflection and Assessment-Vanderbilt Assessment of Leadership in Education. *Webology (ISSN: 1735-188X)*, 18(3).
3. Bhat SK, Prasad KHL. Item analysis and optimizing multiple-choice questions for a viable question bank in ophthalmology: A cross-sectional study. *Indian J Ophthalmol*. 2021 Feb;69(2):343-346. doi: 10.4103/ijo.IJO_1610_20. PMID: 33463588; PMCID: PMC7933874.
4. Boland, R. J., Lester, N. A., & Williams, E. (2010). Writing multiple-choice questions. *Academic Psychiatry*, 34, 310-316.
5. Burud, I., Nagandla, K., & Agarwal, P. (2019). Impact of distractors in item analysis of multiple-choice questions. *International Journal of Research in Medical Sciences*, 7 (4), 1136-1139.
6. Burud, I., Nagandla, K., & Agarwal, P. (2019). Impact of distractors in item analysis of multiple choice questions. *International Journal of Research in Medical Sciences*, 7(4), 1136-1139
7. Elgadal, Amani & Mariod, Abdalbasit. (2021). Item Analysis of Multiple-choice Questions (MCQs): Assessment Tool For Quality Assurance Measures. *Sudan Journal of Medical Sciences*. 16. 10.18502/sjms.v16i3.9695.
8. Ingale AS, Giri PA, Doibale MK (2019). Study on item and test analysis of multiple choice questions amongst undergraduate medical students. *International Journal Of Community Medicine And Public Health*. 2019;4(5):1562-5.
9. Jannah, R., Hidayat, D. N., Husna, N., & Khasbani, I. (2021). An item analysis on multiple-choice questions: a case of a junior high school English try-out test in Indonesia. *Leksika: Jurnal Bahasa, Sastra dan Pengajarannya*, 15(1), 9-17.
10. Juliana D'Sa and Maria Liza (2017) Analysis of Multiple Choice Questions: Item Difficulty, Discrimination Index and Distractor Efficiency. *International Journal of Nursing Education*.
11. Kaipa, R. M. (2021). Multiple choice questions and essay questions in curriculum. *Journal of Applied Research in Higher Education*, 13(1), 16-32.
12. Kheyami, D., Jaradat, A., Al-Shibani, T., & Ali, F. A. (2018). Item Analysis of Multiple Choice Questions at the Department of Paediatrics, Arabian Gulf University, Manama, Bahrain. *Sultan Qaboos University medical journal*, 18(1), e68–e74. <https://doi.org/10.18295/squmj.2018.18.01.011>
13. 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. doi: 10.1016/j.mjafi.2020.11.007. Epub 2021 Feb 2. PMID: 33612937; PMCID: PMC7873707.
14. Loh KY, Elsayed I, Nurjahan M, Roland G(2018). Item difficulty and discrimination index in single best answer MCQ: end of semester examinations at Taylor's clinical school. *Redesigning learning for greater social impact*: Springer; 2018. p. 167-71.
15. Lok Raj, Sharma (2020) Analysis of Difficulty Index, Discrimination Index and Distractor Efficiency of Multiple Choice Questions of Speech Sounds of English. *International Research Journal of MMC (IRJMMC)*, Vol. 2, Issue 1.

16. Maharani, A. V., & Putro, N. H. P. S. (2020). Item analysis of English final semester test. *Indonesian Journal of EFL and Linguistics*, 5(2), 2020
17. Patil R, Palve SB, Vell K, Boratne AV (2017). Evaluation of multiple choice questions by item analysis in a medical college at Pondicherry, India. *International Journal Of Community Medicine And Public Health*. 2017;3(6):1612-6.
18. Rao, C., Kishan Prasad, HL., Sajitha, K., Permi, H., & Shetty J. (2019). Item analysis of multiple choice questions: Assessing an assessment tool in medical students. *International Journal of Educational and Psychological Researches*, 2 (4), 201-204.
19. Shahid, R., Farooq, Q., & Iqbal, R. (2019). Item analysis of multiple choice questions of ophthalmology at Rawalpindi Medical University, Rawalpindi, Pakistan. *Rawal Medical Journal*, 44(1), 192-192.
20. Sharma, L. R. (2021). Analysis of difficulty index, discrimination index and distractor efficiency of multiple choice questions of speech sounds of English. *International Research Journal of MMC (IRJMMC)*, 2(1), 15-28.
21. Stringer, J. K., Santen, S. A., Lee, E., Rawls, M., Bailey, J., Richards, A., ... & Biskobing, D. (2021). Examining Bloom's taxonomy in multiple choice questions: students' approach to questions. *Medical Science Educator*, 31(4), 1311-1317.