RESEARCH ARTICLE

ANATOMY MULTIPLE CHOICE QUESTIONS ITEM ANALYSIS: DISTRACTOR EFFICIENCY ASSOCIATION WITH DIFFICULTY AND DISCRIMINATION INDICES

Khadija Qamar[1], Muhammad Alamgir Khan[2], Sayed Nusrat Raza[3], Aliya Hisam[4], Rehana Khadim[5]

[1] Department of Anatomy , Army Medical College, National University of Medical Sciences
[2] Department of Physiology, Army Medical College, National University of Medical Sciences
[3] Principal & Dean, Army Medical College, National University of Medical Sciences
[4] Department of Community Medicine ,Army Medical College, National University of Medical Sciences
[5] Army Medical College, National University of Medical Sciences

ABSTRACT

Objective: *To determine the association of distractor efficiency with difficulty and discriminatory indices in Anatomy multiple choice questions (MCQs).*

Design: *Cross-sectional.*

Place and Duration of Study: *Department of Anatomy, Army Medical College, National University of Medical Sciences Rawalpindi from January 2020 till July 2020.*

Materials and Methods: *About 215 MCQs of Anatomy of both MBBS and BDS classes through convenience sampling were selected. The MCQs were obtained through optical marks recognition data from the examination branch along with variables of interest, i.e., distractor efficiency, difficulty index and discrimination index. Data were analyzed in STATA Version 14 and SPSS version 26.*

Results: *The mean distractor efficiency was 64.96 ± 34.28, the mean difficulty index was 65.22±22.53 and the mean discrimination index was 0.30±0.13. Distractor efficiency has a significantly strong negative association with difficulty index (r = -0.73, p=0.001) while a significantly weak positive association with discrimination index (r= 0.20, p=0.002).*

Conclusion: *The Anatomy MCQs item analysis showed that distractor efficiency is negatively correlated with difficulty index but positively correlated with discrimination index.*

Keywords:

Assessment, Distractor efficiency, Difficulty index, Discrimination index, Multiple Choice Questions

INTRODUCTION

Evaluation or assessment is a step wise process of forming an opinion about the quality and depth of student performance and accomplishment, then concluding regarding the learning process at the end.[1] Assessment drives learning emphasizes the central role of assessment in any form of education and particularly in high stake medical education assessments. Multiple choice questions (MCQs) items are the most common method of assessing the knowledge capabilities of undergraduate students in medical colleges. Framing meaningful MCQs is a time consuming and challenging process. The correctly formed MCQs result in unbiased analysis that measures understanding, knowledge, application, and assessment.[2] MCQs to be used must be of good quality, and they need to be tested for that standard or quality. Interpretation of post-exam item analysis report and modification or improvement of MCQs accordingly is an essential prerequisite to maintain good quality MCQ bank as required by the regulatory bodies like Pakistan Medical Council (PMC). Item analysis[3] is an essential step in the development of any assessment strategy.[4] The phase helps us identify an item that is either too difficult or too easy for the examinee. This process also helps in detecting items that fail to discriminate between skilled and unskilled examinees. Item analysis also gives a view of the process of instructions. If correctly sampled,

Correspondence:
Dr. Rehana Khadim
Managing Editor PAFMJ, Army Medical College, National University of Medical Sciences
E.mail: dr.rehana.butt@gmail.com

the items will inform whether the subject/topic/concept has been learnt and understood by the learner. Proper assessments will decipher the grey areas, misconceptions and low clarity zones that require the attention of the learner and facilitator both.

In item analysis, the process checks the effectiveness of test items by the score of the exam and sorts the results by score. Based on student results, the measurement of difficulty and discrimination indices and the correlation of marks will provide proof of validity. The complexity of the test item and its discrimination power (DP) may provide supporting evidence of the accuracy of the examinations.[5]

The difficulty index is the percentage of participants who correctly attempts the exam item (e.g. single question in the best response form MCQ paper).[6] Its calculation is simply by dividing the participants who have passed the exam item by the participants who have not cleared the exam items.[7] A simple item may have a high discrimination index (e.g., 0.9) while a hard item may have a low discrimination index (e.g., 0.1). The degree to which success or failure on a test item indicates possession of the ability being measured is referred to as the discriminatory index of a test item. It determines how discriminating an item is among examinees in terms of the function or ability it measures. This number might be anywhere between 0.0 and 1.00. The higher the value, the greater the item's discrimination.[8] The discrimination power has the capability to differentiate between a low performing student and a high performing student. The DP is the ability of a test item to distinguish between high and low performers. For DP calculation, firstly the students are ranked in order of their scores/marks achieved in the examination. Then the students in the uppermost third and bottom-most third are calculated. Then the percentage of students correctly attempting the MCQs item in the bottommost third is subtracted from the percentage of students in the. uppermost third. A positive DP is desired. A negative DP needs to be assessed and corrected.[7,8]

A MCQs item analysis showing a low discrimination index may have many reasons: that is it may be out of the syllabus; the teaching methodology was flawed or ineffective; the construct of the item is not proper; students have not understood the topic properly.[9, 10] An error in the MCQs can be detected by a negative value: it means that the students identified as low performers are answering more accurately than the students who are high performers: this means that the MCQs item is flawed in the construct, examination misconduct was there or the answer was incorrect. Therefore, a negative discrimination index needs evaluation and improvement. The MCQs construct shall aim to avoid non-functional distractors completely and add functional distractors, the ones that are incorrect but shall be able to differentiate between low and high achievers.[11]

Distractors had been extensively investigated in terms of guessing probability, but their impact on other item analysis parameters like difficulty and discriminatory indices generally remained out of focus. Hence, this study was planned to determine the association of distractor efficiency with difficulty and discriminatory indices in Anatomy MCQs at Army Medical College, Rawalpindi.

METHODOLOGY

This cross-sectional study was carried out at the department of Anatomy, Army Medical College, National University of Medical Sciences. The study commenced after obtaining permission from the Institutional Ethical Review Committee. The sample size was calculated by using the software G-Power version 3.1.9.4, Considering the values of effect size as 0.20, alpha error probability as 0.05 and power of the test as 80% as the sample size of 191 was calculated. However, we took a sample of 215 MCQs of Anatomy of both MBBS and BDS classes through convenience sampling. Optical marks recognition (OMR) data related to the MCQs were obtained from the examination branch, and variables of interest, i.e., distractor efficiency, difficulty index and discrimination index, were analyzed.

Distractor efficiency was calculated based upon a number of non-functional distractors (NFDs). Any distractor attempted by less than 5% of the students was declared as NFD. Zero NFD means 100% distractor efficiency, 1 NFD means 66.67% distractor efficiency, 2 NFD means 33.33% distractor efficiency and 3 NFD means 0% distractor efficiency.

Based on the difficulty index, the MCQs were divided into three categories. MCQs with less than 30% difficulty index were classified as 'hard', those from 30 to 70% difficulty index were classified as 'moderate', and the MCQs with greater than 70% difficulty index were classified as 'easy' MCQs.

Similarly, the MCQs were divided into three categories based on the discrimination index. MCQs having

discrimination index less than 0.2 were classified as poor, those from 0.2 to 0.39 as good and the MCQs with discrimination index over 0.4 were classified as excellent.

The data were entered in STATA Version 14 and SPSS version 26 for analysis. For continuous variables, mean, and the standard deviation was calculated while for categorical variables, frequency and their percentages were calculated. The correlation between numerical variables was calculated using the spearman correlation coefficient as the data was not normally distributed. Whereas the association between categorical variables was calculated using the Chi-Square test. The p-value of less than or equal to 0.05 was considered significant.

RESULTS

Descriptive analysis of 215 Anatomy MCQs is shown in figure 1. Pie Charts in the figure (Ia, Ib, Ic) shows the frequency and percentage of the MCQs in each category of distractor efficiency, difficulty index and discrimination index. Figure Id shows the mean and standard deviation of all three variables.

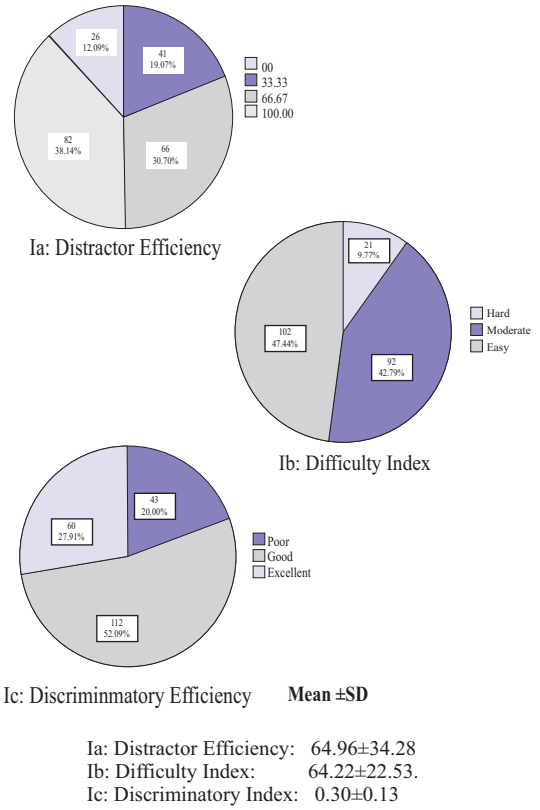

Fig 1: Descriptive analysis of distractor efficiency, difficulty index and discriminatory index showing frequency, percentage, mean and standard deviation (SD)

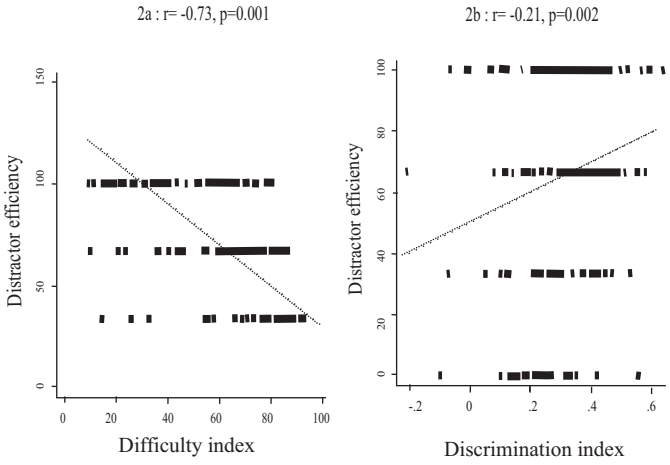

Figure 2: Correlation of distractor efficiency with difficulty and discrimination indices

The scatter diagram in figure 2 shows the correlation of distractor efficiency with difficulty and discrimination indices along with values of correlation coefficient and alpha error (p-value). Figure 2a shows a significant negative correlation between distractor efficiency and difficulty index ($p<0.001$) while a significant positive correlation between distractor efficiency and discrimination index ($p=0.002$). However, the strength of the correlation is high with difficulty index but weak with discrimination index as shown by the 'R-values'.

Frequency/percentage comparison of MCQs in each category of distractor efficiency and difficulty index is shown as cross-tabulation in table 1. It is evident from the table that out of the 215 MCQS, as there is an increase in distractor efficiency, the difficulty index increases. The distractor efficiency of 33% and 66.67% have mostly the moderate and easy difficulty index, which means that there is at least 1-2 distractor with 100% distractor efficiency.

The table I shows that frequency/percentage of MCQs in 'hard' category of difficulty index increases with the rise in distractor efficiency whereas that in 'easy' category decreases and the association between the two variables is statistically significant ($0<0.001$).

Table II shows the frequency/percentage comparison of MCQs in each category of distractor efficiency and discrimination index. The frequency/percentage of MCQs in 'excellent' category of discrimination index increases with the rise in distractor efficiency whereas that in 'poor' category decreases and the association between the two variables is statistically significant ($p<0.001$).

Table I: Association of distractor efficiency with a difficulty index

		Difficulty index			p-value
		Hard	Moderate	Easy	
Distractor efficiency	0%	0	0	26(100%)	<0.001
	33.33%	2(4.88%)	6(14.63%)	33(80.49%)	
	66.67%	3(4.55%)	29(43.94%)	34(51.52%)	
	100.00%	16(19.51%)	57(69.51%)	9(10.98%)	

Table II: Association of distractor efficiency with discrimination index

		Difficulty index			p-value
		Poor	Good	Excellent	
Distractor efficiency	0%	11 (42.31%)	13 (50.00%)	2 (7.69%)	<0.001
	33.33%	6 (14.63%)	26 (63.41%)	9 (21.95%)	
	66.67%	13 (19.70%)	32 (48.48%)	21 (31.82%)	
	100.00%	13 (15.85%)	41 (50.00%)	28 (34.15%)	

The total Anatomy MCQs item analysis has shown the mean, standard deviation of distractor efficiency to be 64+34.28, difficulty index of 65.22+22.53 and discrimination index of 0.3+0.13. There are 38.14% of the MCQs having three functional distractors with the majority of them having a moderate difficulty index (42.79%) and a discrimination index of 52.09%. A strong significant negative association (r= -0.73, p=<0.001) is found between distractor efficiency and difficulty index while a weak positive significant association (r= 0.21, p<0.002) between distractor efficiency and discrimination index is established. In the Anatomy MCQs item analysis, there was a large number of MCQs with having three functional distractors of moderate difficulty index (69.51%). While the MCQs having one and two distractors were mostly of the type having moderate (43.94%) and easy difficulty index ((80.49%) respectively.

DISCUSSION

In the educational institute, the MCQ quality can be evaluated on the distractor efficiency, difficulty index and discrimination index.[12] The distractor efficiency, difficulty index and discrimination index are among the instruments to check whether the MCQs are well constructed or not. Distractor efficiency analyses the quality of distractors and is closely associated with difficulty and discrimination indices. A distractor used by less than 5% of students is not a significant distractor and should be either replaced or corrected as it affects the overall quality of the question.

In a study conducted in Saudi Arabia Nursing Institute, the mean distractor efficiency was 74.3% which is more than the mean distractor efficiency of this study that is 64.96+ 34.28. In the same study, the difficulty index and discrimination index had a negative correlation (r = -0.721; p<0.01) while in this study it was also a negative one that is - 0.08 but increased in comparison.[13]

The distractor efficiency of difficult items in this study is 100%, which was expected, as difficult items would require much guesswork on the part of the student, thereby using all the distractors. We observed that items having one NF-D had excellent discriminating ability (difficulty index = 0.427) as compared to items with all four functioning distracters (difficulty index = 0.351). This compares well with other studies favouring better discrimination by three distracters as compared to four.[14]

Distractor construction is often a difficult task and yet an essential element of an MCQ construct. A distractor selected by <5% of the graduate is categorized as a poor distractor.[6] It is quite challenging to construct an MCQ with more than three plausible distractors, and so the fillers are added.[15] In one study, 39% of the MCQs were having two plausible distractors, indicating difficulty in the construction of three plausible distractors. Through item analysis, the instructor shall be able to remove non-plausible distractor and replace it with a more appropriate one.

In a study conducted in Pakistan, many MCQs were having difficulty index of 81% and discrimination index of 83% in comparison to this study which has shown 47.44% of easy difficulty index and 28% of excellent discrimination index.[16,17] The item analysis of 48 MCQS showed the mean difficulty index of 67.5 and mean discrimination of 0.44 as compared to the 65.22 and 0.3 respectively.[12]

The quality of MCQs has a significant impact on the test analysis. About 40 MCQs analysis showed a mean discrimination index of 0.22, which is lower than in this study that is 0.3+0.13.[18] As we have not randomly selected the MCQ for analysis, there might be a chance of having sampling bias. So the association between distractor efficiency and discrimination index must be re-evaluated in further studies with larger sample sizes and random sampling. The evaluation of student's cognitive knowledge through MCQ shall be based on information of the subject rather than recall only. In another study, the discrimination index was 0.26 which is again less than this study that is 0.3+0.13.[19]

Framing concise MCQs is a time-consuming and challenging process. It is said that appropriately

constructed MCQs result in objective testing that can measure knowledge, comprehension, application, analysis, and evaluation.[2] The MCQ medium is English which is not the native language of the student is so the evaluation might be affected by the language bias/reading ability. It is not possible to construct medicine MCQ in the Urdu language, but if in future they are constructed, the language validity of the MCQ can be checked through item analysis in all the MCQ pools.

CONCLUSION

Distractor efficiency is negatively correlated with difficulty index but positively correlated with discrimination index.

While constructing MCQs, this fact needs to be kept in mind because the addition of non-functional distractors, which is not uncommon, will reduce the quality of MCQ by not only decreasing the distractor efficiency but adversely affecting other item analysis variables like difficulty and discrimination indices.

ETHICALAPPROVAL

The Ethical Review Committee approval was taken before the conduction of the study.

CONFLICT OF INTEREST

The authors declared no conflict of interest.

SOURCE FUNDING

NIL

REFERENCES

1. Dent J, Harden RM, Hunt D. A practical guide for medical teachers: Elsevier health sciences; 2017.
2. Schuwirth LW, van der Vleuten CP. Different written assessment methods: what can be said about their strengths and weaknesses? Medical education. 2004;38(9):974-9. 10.1111/j.1365-2929.2004.01916.x
3. Wass V, Van der Vleuten C, Shatzer J, Jones R. Assessment of clinical competence. Lancet. 2001;357(9260):945-9. 10.1016/s0140-6736(00)04221-5
4. Ul Islam Z, Usmani A. Psychometric analysis of Anatomy MCQs in Modular examination. Pakistan journal of medical sciences. 2017;33(5):1138
5. Daugherty R. Assessment for Learning: Research and Policy in the (Dis)United Kingdom: Springer; 2011 August 2020]. Available from: https://link.springer.com/ chapter/10.1007/978-94-007-0729-0_2#citeas.
6. F hMaJ. Analysis of One-Best MCQs: the Difficulty Index, Discrimination Index and Distractor Efficiency. JPMA. 2012;62(2):142-7
7. Chappuis S CCaSR. Balanced assessment systems: Leadership, quality, and the role of classroom assessment 2020 August 2020]. Available from: https://resources.corwin.com/chappuisbalancedassessment.
8. Boopathiraj C, Chellamani K. Analysis of test items on difficulty level and discrimination index in the test for research in education. International journal of social science & interdisciplinary research. 2013;2(2):189-93
9. Baysinger MA, Scherer KT, LeBreton JM. Exploring the disruptive effects of psychopathy and aggression on group processes and group effectiveness. The Journal of applied psychology. 2014;99(1):48-65. 10.1037/a0034317
10. DeSimone JA, James LR. An item analysis of the Conditional Reasoning Test of Aggression. The Journal of applied psychology. 2015;100(6):1872-86. 10.1037/apl0000026
11. Testa S, Toscano A, Rosato R. Distractor Efficiency in an Item Pool for a Statistics Classroom Exam: Assessing Its Relation With Item Cognitive Level Classified According to Bloom's Taxonomy. Frontiers in Psychology. 2018;9(1585). 10.3389/fpsyg.2018.01585
12. D'Sa JL, Visbal-Dionaldo ML. Analysis of multiple choice questions: item difficulty, discrimination index and distractor efficiency. International Journal of Nursing Education. 2017;9(3):109-14
13. Ahmed IAM. Role of Multiple Choice Questions in Assessment of Competency of Knowledge in Anatomy. Health Sciences. 2020;9(6):72-8
14. Tarrant M, Ware J, Mohammed AM. An assessment of functioning and non-functioning distractors in multiple-choice questions: a descriptive analysis. BMC medical education. 2009;9(1):40
15. Haladyna TM, Downing SM. How many options is enough for a multiple-choice test item? Educational and psychological measurement. 1993;53(4):999-1010

16. 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. Annals of PIMS-Shaheed Zulfiqar Ali Bhutto Medical University. 2017;13(4):310-5

17. Mahjabeen W AS, Zafar T, Butt R, Konain S and Rizvi M. Difficulty Index, Discrimination Index and Distractor Efficiency in Multiple Choice Questions. Annals of PIMS. 2017;13(4):310-5

18. Charania D. Item Analysis of Multiple Choice Questions Given to First Year Medical Students–Concept Building and MCQS. Int J Research in Humanities. 2015:3(9); 18-29

19. Melovitz Vasan CA, DeFouw DO, Holland BK, Vasan NS. Analysis of testing with multiple choice versus open-ended questions: Outcome-based observations in an anatomy course. Anatomical sciences education. 2018;11(3):254-61. 10.1002/ase.1739