Three Option Multiple Choice Questions had the Least Non-functioning Distractors: Analysis of 1855 MCQs in First Year Competency Based Medical Program at Tanta Faculty of Medicine, Egypt

Magdy Balaha (1), Manal El Bermawy (2) and Eslam El-Hawary (3)

Abstract

Introduction: The study aims to evaluate the efficiency of distractors by analyzing the items used in 18 courses along the first year of the CBMBP curriculum. The study is among few large studies discussing distractor efficiency in the Middle East. It is the largest in Egypt and Arabic area. **Settings**: Tanta Faculty of Medicine, Egypt from December 2016 until July 2017. This study was conducted on the CBMBP exams after the end of the semesters of the first year. **The study design** was repeated cross-sectional study. 16 courses, 28 tests and 1855 MCQs were evaluated. **Methods:** The MCQs were analyzed using Difficulty Index (Dif I), Discrimination Index (Dis I), using the point biserial (pbs) and analysis of distractor efficiency (DE) by calculation the non-functioning distractors (NFD). **Results**: The mean difficulty index (Diff I) was 0.71, 46% of the questions in the easy category while the difficult questions represented 11%. The mean discrimination index (Disc I) was 0.27 with 15% of the questions having poor discriminatory value. The presence of NFDs as following; 0 NFDs in 30%, 1 NFDs n 40 %, 2 NFDs in 23% and those with >= 3 NFDs in 7%. Distractor efficiency (DE%) was 56.5%. The number of NFDs increased with increasing number of options. **Conclusion**: The number of NFDs increased with increasing number of options. The best values of NFDs were found in the three options MCQs. The 3 option MCQs had the best number of good distractors with 0 NFDs, hence they have the reasonable difficulty and best discrimination.

Key words (Distractor efficiency, nonfunctioning distractors, Item difficulty, Discrimination)

Authors

Magdy Balaha (1): Professor of Gynecology (**Corresponding Author**)

E mail: abcd221221@gmail.com; Magdy.balaha@med.tanta.edu.eg

Manal El Bermawy (2): Professor of Anatomy

Email: manalelbermawy1@gmail.com; Manal.elbermawy@med.tanta.edu.eg

Eslam El-Hawary (3): Lecturer of Pediatrics

Email: manalelbermawy1@gmail.com; Manal.elbermawy@med.tanta.edu.eg

Tanta University, Tanta Faculty of Medicine, Egypt

1 INTRODUCTION

Our medical school; Faculty of Medicine (FOM), Tanta University (TU), Tanta, Egypt, was established in 1963 with the first batch graduated at 1969. Since then, the traditional Flexnerian curriculum has been the only curriculum adopted by the FOMTU for the subsequent 48 batches graduated until now. In 2016, the Faculty started an innovative change in the curriculum towards a competency-based one.
Currently, the Faculty is implementing both the traditional Flexnerian curriculum for the main student batch parallel to a competency-based curriculum (CBMBP) applied for a limited number of students. As regard the evaluation process, the final written exam in the traditional curriculum has been depending on classical long and short essay questions until 2016, when multiple-choice questions (MCQs) were adopted and made up to 50% of the marks. On the other hand, CBMBP curriculum adopted lots of continuous assessment methods; including all forms of non-MCQ written assessments, but the midterm and final written exams were all made up of MCQs.
For the evaluation process to be comprehensive, different approaches including written assignments, oral presentations, essay, short-answer and multiple-choice questions (MCQs) need to be integrated [1]. MCQs retain a favorable position compared to the rest of assessment tools. This can be attributed to the fact that MCQs are easily and reliably scored. Furthermore, they provide a wider range of material sampling along the entire course work, which eventually increases the test validity. MCQs are considered more reliable in evaluation of the students' knowledge than short answer questions [2].

There are different types of MCQs, but the single best answer or A-type remains by far the most common. In each question (item), a stem is followed by a number of answer options. The options include a single correct answer (key) and the remaining are incorrect options or distractors. The pre-validation stage, which evaluates the MCQs depending on fulfilment of the quality parameters, and the post-validation stage where item analysis plays an essential role [3], [4].

Item analysis is a process that examines students' responses to individual test items to evaluate the quality of these items and the quality of the test as a whole. It offers a great help in improving the quality of items, that may be used again in subsequent tests. It can help sort out poor items, which need improvement or even deletion. Feedback also improves the skill of item construction and provides a clue about which parts of the course that needs greater emphasis or clarity. [5], [6].

Typically, in a test analysis, three values are computed; a difficulty level, a discrimination index and distractors efficiency. The difficulty index refers to the difficulty of an item for students to identify the correct choice. The discrimination index indicates how effectively the item discriminates strong students from week ones. Distractors efficiency describes the ability of each of the provided distractors to distract some students from the key answer, and hence they are considered as functioning distractors [7].

Review of literature revealed that many international and regional papers were published concerning the significance of post-validation concepts as well as their acceptable values. Nevertheless, there were very few Egyptian studies in this domain. There is only one publication concerned with board examination in Family Medicine (Khafagy et al 2016) [8].

To the knowledge of the authors, this study is among few large studies of item analysis of MCQs in the Middle East. It is the largest in Egypt and Arabic area.

The aim of the study was to evaluate the efficiency of distractors by analyzing the items used in 18 courses along the first year of the CBMBP curriculum and studying difficulty index, discrimination index and distractor efficiency.

2 METHODS

2.1 SETTINGS:

The present study was conducted in Tanta Faculty of Medicine (FOM), Tanta University (TU), Tanta, Egypt from December 2016 until July 2017. The College Dean approved this study and allowed access to the examination data in the exam control unit. The identities of the students taking the examination were kept anonymous and confidential.

2.2 POPULATION:

This study was conducted on the CBMBP exams after the end of the semesters of the first year. There were four time slots of exams; two times in each semester. (mid-term as well as final written exam). Each exams was taken by legible students of 23.

2.3 DESIGN:

Repeated cross-sectional study.

2.4 PROCEDURE:

In the new CBMBP the only method of final written exams were MCQs with single best response. Two, three, four, and five options were allowed, with the majority of MCQs having 4 options.

Pre-validation of the test paper was done by the course coordinator as well as the semester coordinator. The pre-validation included checking test validity as well as MCQ construction. All items were uploaded to a question bank, from which all exams are selected shortly before the time of examination.

Each correct response was awarded one mark, a wrong response was given 0 marks with no negative marking. The students' answers were written down in premade answer papers that were subsequently scanned by the OMR machine to be computed as a spreadsheet in the OMR program. Calculations were done both automatically by the OMR machine (Remark office OMR V.10 soft program) and manually on the spreadsheets for cross validation.

Post validation testing of the exam paper was done by item analysis. The scores of all the students were arranged in order of merit. The upper one-fourth (highest quarter) students were considered as high achievers while the lower fourth (lowest quarter) were considered as low achievers. The MCQs were analyzed using various indices like Difficulty Index (Dif I), Discrimination Index (Dis I), and analysis of distractor efficiency (DE) by calculation the non-functioning distractors (NFD).

2.5 INTERPRETATION:

2.5.1. Difficulty Index (Dif I or p) describes the percentage of students who selected the correct response and ranges between 0 and 100%. The higher the value of Dif I, the easier the item is and vice-versa. There are many published cut-off values depending on some factors as the academic level of students (novice or graduating exams) and the number of question options (2, 3, 4, and 5 options). In this study, we specified the Dif I cut-off values to be; <0.3 difficult, 0.3-0.8 average and > 0.8 easy because the study population were beginner first year students and also because we did not use a fixed number of options in all the questions. [9], [10]

2.5.2. Discrimination Index (item effectiveness - Dis I) indicates how effectively the question can sort out (discriminate) the students who actually know the material from those who do not. There are many indices to express the discrimination. In this study, the point biserial (pbs) was used to express the Disc Index. The pbs ranges from -1 to +1. Dis I of 1 is considered as ideal, which can efficiently discriminate between high and low achievers. An item having a pbs greater than or equal to 0.18 is considered to have excellent discriminative and between 0 - 0.18 with acceptable discriminative. An item having negative discrimination < 0.0 (negative Dis I) has poor discriminative power. [9], [10], [11]

2.5.3. Distractor Efficiency analysis (DE or NFD distribution): Any item contains a stem and two, three, four or five options;

including one correct (key) and other incorrect alternatives or distractors. Functioning distractor (FD) means that the frequency of this distractor choice is >5% of students; in other words, functional or effective distractors are those selected by 5% or more participants. In four option items, if the item contains three, two, one, or nil NFDs, its distractor efficiency (DE) would be 0, 33.3%, 66.6% and 100% respectively. In five option items, if the item contains four, three, two, one, or nil NFDs, its distractor efficiency (DE) would be 0%, 25%, 50%, 75% or 100% respectively. In three option items, if the item contains two, one, or nil NFDs, its distractor efficiency (DE) would be 0%, 50%, or 100% respectively. [12], [13], [14], [15]

2.6 STATISTICAL ANALYSIS:

In the formed spreadsheets of all exams, item analysis was done as a cross validation for machine calculation. Various indices like difficult index (DIF I), discrimination index (DI) and Distracter analysis were calculated. In addition, counting and simple analysis was done. Further data analysis was performed using the Statistical Package for the Social Sciences (SPSS), Version 25.0 (IBM Corp., New York, USA). P value for significance were considered at 0.05 and 0.01.

3 RESULTS

23 students gave 28 tests throughout the first year of the CBMBP, consisting of 1855 MCQ items. **As seen in Table 1**, the number of MCQs in mid semester 1 was 330, with a mean difficulty Index (Dif I) of 0.76 ± 0.02 %, which is considered average as it lies between 0.3-0.8. The same applies to the mean Dif I of the final semester 1, mid semester 2 and final semester 2 exams, which were 0.64 ± 0.01, 0.74 ± 0.02 and 0.69 ± 0.012 respectively.

The Discrimination Index (Dis I) on the other hand was widely variable from course to course and from semester to another. For example, the Cell biology test in med semester 2 had the lowest mean Dis I (0.11 ± 0.02), while the Human body test in final semester 1 had the highest Dis I (0.60 ± 0.03). Despite the variability between different courses, none of them had negative Dis I, and the greater part had Dis I more than 0.18.

Distractor Efficiency analysis (DE or NFD distribution), **table 2 & Fig 1a,b** compares the Distribution of NFD per the number of MCQ options. The overall analysis of MCQs with three options revealed the absence of NFD in 39% of items, while 46% of items had 1 NFD, 15% of items had 2 NFD. In four option MCQs, the absence of NFD was noted in 24.7% of items, while 37.3% of items had 1 NFD, 26.8% of items had 2 NFD and 11.2% of items had 3 NFD. (the more the number of options the more the chance to have non-functioning distractors).

By comparing the mean difficulty and discrimination indices with the number of non-functioning distractors **(table 3, Fig 2a,b)** the authors found that there was a strong correlation between them. In the items of semester one, items with complete absence of NFD had mean Dif. I. of 0.51 and a mean Dis. I of 0.37. In the items of semester two, items with complete absence of NFD had mean Dif. I. of 0.52 and a mean Dis. I of 0.24.

Looking at correlation the number of options in each question to the number of NFD **(table 3),** the authors found that there was a strong positive correlation between the number of options and the number of NFD. There was also a positive correlation between the number of NFD and the corrected difficulty index, while the correlation was negative between the number of NFD and the discrimination index.

Table 1: Mean difficulty index, discrimination index and distractor efficiency of all CBMBP first year examinations at Tanta College of Medicine, Tanta, Egypt (28 courses, N = 1855 MCQs)

Course	N	Dif.I %	Dis.I	DE %
Mid Semester 1		**Mean ± SE**		
Human body	80	0.74 ± 0.02	0.19 ± 0.02	54.5 ± 3.34
Basic English	50	0.85 ± 0.04	0.12 ± 0.02	26.9 ± 3.84
English Terminology	50	0.78 ± 0.03	0.37 ± 0.03	48.7 ± 5.03
Patient Safety	50	0.78 ± 0.04	0.11 ± 0.02	34.9 ± 4.47
Biostatistics	50	0.84 ± 0.04	0.08 ± 0.02	27.5 ± 4.31
Medical Informatics	50	0.58 ± 0.04	0.36 ± 0.04	65.3 ± 3.83
Total mid S1 MCQs	**330**	**0.76 ± 0.02**	**0.23 ± 0.01**	**43.9 ± 1.84**
Final Semester 1				
Human body	80	0.66 ± 0.02	0.60 ± 0.03	75.24 ± 3.23
Basic English	65	0.63 ± 0.04	0.19 ± 0.02	66.32 ± 4.18
English Terminology	100	0.74 ± 0.02	0.41 ± 0.02	63.89 ± 3.09
Patient Safety	80	0.54 ± 0.03	0.31 ± 0.03	74.84 ± 2.77
Biostatistics	100	0.64 ± 0.03	0.39 ± 0.03	70.79 ± 3.08
Medical Informatics	60	0.64 ± 0.03	0.37 ± 0.03	80.00 ± 3.80
Management skills	50	0.64 ± 0.05	0.29 ± 0.03	67.72 ± 3.85
Professionalism 1	70	0.72 ± 0.04	0.18 ± 0.02	54.71 ± 3.70
Education skills	60	0.53 ± 0.04	0.36 ± 0.04	75.92 ± 3.96
Total Final S1 MCQs	**665**	**0.64 ± 0.01**	**0.34 ± 0.01**	**69.71 ± 1.18**
Mid Semester 2				
Cell biology	50	0.84 ± 0.04	0.11 ± 0.02	50.40 ± 4.7
Biochemistry	50	0.74 ± 0.04	0.38 ± 0.04	52.92 ± 3.92
Genetics	50	0.73 ± 0.05	0.21 ± 0.02	48.04 ± 4.5
Histology	50	0.82 ± 0.03	0.25 ± 0.02	54.86 ± 5.4
Anatomy	60	0.68 ± 0.04	0.22 ± 0.02	57.77 ± 4.03
Physiology	60	0.68 ± 0.04	0.20 ± 0.03	57.27 ± 3.17
Total mid S2 MCQs	**320**	**0.74 ± 0.02**	**0.23 ± 0.01**	**53.79 ± 1.74**
Final Semester 2				
Cell biology	80	0.72 ± 0.03	0.28 ± 0.02	63.9 ± 3.8
Biochemistry	80	0.83 ± 0.03	0.16 ± 0.02	44.3 ± 3.5
Genetics	80	0.82 ± 0.03	0.19 ± 0.02	50.4 ± 3.7
Histology	80	0.64 ± 0.04	0.23 ± 0.02	64.2 ± 3.74
Anatomy	90	0.63 ± 0.03	0.28 ± 0.02	73.57 ± 2.57
Physiology	90	0.59 ± 0.03	0.19 ± 0.02	67.13 ± 3.14
Interviewing skills I	40	0.82 ± 0.03	0.21 ± 0.02	56.43 ± 4.8
Total final S2 MCQs	**540**	**0.69 ± 0.012**	**0.22 ± 0.008**	**56.2 ± 6.57**
Average		***0.7 ± 0.006***	***0.26 ± 0.005***	***59.7 ± 0.76***

Exam. = examination; SE = standard error; Dif.I = difficulty index; Dis.I = discrimination index; DE = distractor efficiency, S= Semester, N = number of MCQs

IJSER

Table 2: Distribution of Non-functioning distractors per the number of MCQ options in all CBMBP first year courses' examinations at Tanta College of Medicine, Tanta, Egypt (2 semesters, 28 courses, N = 1855 MCQs)

MCQ Options	N (%)	Distribution of NFD per the number of MCQ					Total
		0	1	2	3	4	
All S1							
3	**n (%)**	**163 (36)**	**221 (48)**	**73 (16)**	**1 (0)**	**0 (0)**	458 (100)
4	**n (%)**	**163 (31)**	**184 (35)**	**121 (23)**	**57 (11)**	**0 (0)**	525 (100)
5	n (%)	2 (17)	5 (42)	2 (17)	3 (25)	0 (0)	12 (100)
All S2							
2	n (%)	2 (50)	2 (50)	0 (0)	0 (0)	0 (0)	4 (100)
3	**n (%)**	**97 (46)**	**87 (41)**	**27 (13)**	**0 (0)**	**0 (0)**	211 (100)
4	**n (%)**	**125 (19.5)**	**251 (39.2)**	**191 (29.8)**	**74 (11.5)**	**0 (0)**	641 (100)
5	n (%)	1 (25)	2 (50)	1 (25)	0 (0)	0 (0)	4 (100)
Overall							
3	n (%)	260 (39)	308 (46)	100 (15)	1 (0.1)	0 (0)	669 (100)
4	n (%)	288 (24.7)	435 (37.3)	312 (26.8)	131 (11.2)	0 (0)	1166 (100)

NFDs = non-functioning distractors; Dif.I = difficulty index; Dis.I = discrimination index; S= Semester

Table 3: Mean difficulty and discrimination per number of Non-functioning distractors of all CBMBP first year examinations at Tanta College of Medicine, Tanta, Egypt (2 semesters, 28 courses, N = 1855 MCQs)

Semester Parameter	Distribution of items as per their NFDs					Total	Corr.
	0	1	2	3	4		
Mid S1 n (%)	**52 (15.8)**	**136 (41.2)**	**96 (29)**	**44 (13)**	**2 (0.6)**	**330 (100)**	
Mean Dif.I	0.47	0.71	0.87	0.99	1	0.76	*0.593***
Mean Dis.I	0.33	0.23	0.1	0.002	0.0	0.18	*-0.237***
Final S1 n (%)	**274 (41.2)**	**271 (40.8)**	**103 (15.5)**	**16 (2.4)**	**1 (0.2**	**665 (100)**	
Mean Dif.I	0.51	0.67	0.86	0.98	-	0.64	*0.462***
Mean Dis.I	0.38	0.35	0.26	0.05	-	0.34	*-0.061*
Mid S2 n (%)	**60 (18.8)**	**142 (44.4)**	**85 (26.6)**	**33 1 (10.3)**	**0 (0)**	**320 (100)**	
Mean Dif.	**0.58**	**0.67**	**0.87**	**0.99**	-	**0.74**	*0.459***
Mean Dis.I	**0.24**	**0.26**	**0.21**	**-0.18**	-	**0.23**	*-0.142**
Final S2 n (%)	**165 (30.6)**	**198 (36.7)**	**135 (25)**	**42 (7.8)**	**0 (0)**	**540 (100)**	
Mean Dif.I	**0.51**	**0.69**	**0.86**	**0.99**	-	**0.71**	*0.564***
Mean Dis.I	**0.24**	**0.23**	**0.17**	**0.23**	-	**0.22**	*-0.055*

NFDs =non-functioning distractors; Dif. I = difficulty index; Dis.I = discrimination index;, S= Semester. ##Corr. = Correlation between non-functioning distractors (NFDs) and the Dif. I, Dis. I in in the different MCQs examinations

IJSER

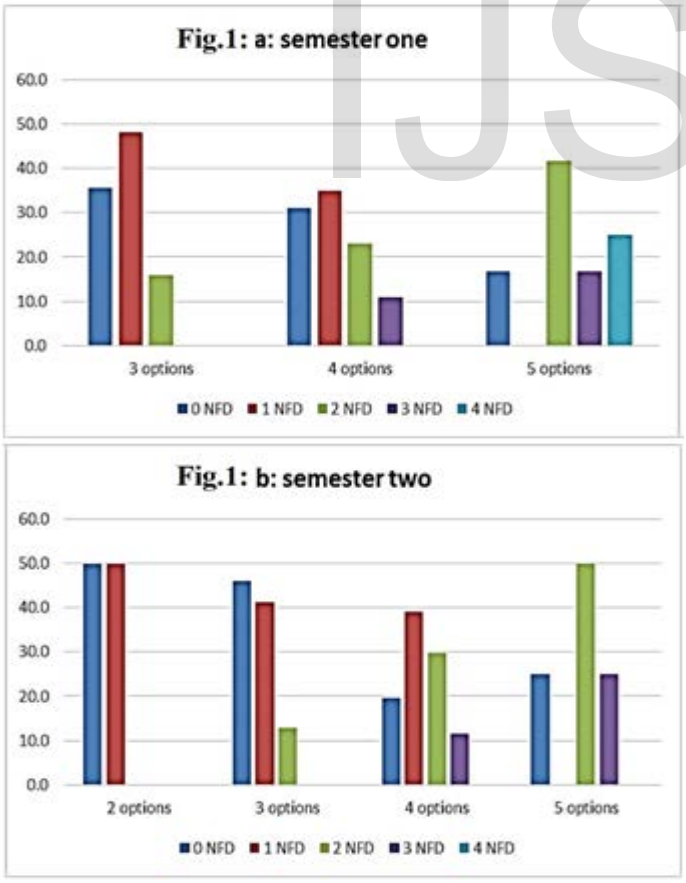

Fig. 1 (a, b): Distribution of the percentage of NFD per number of options in both semesters

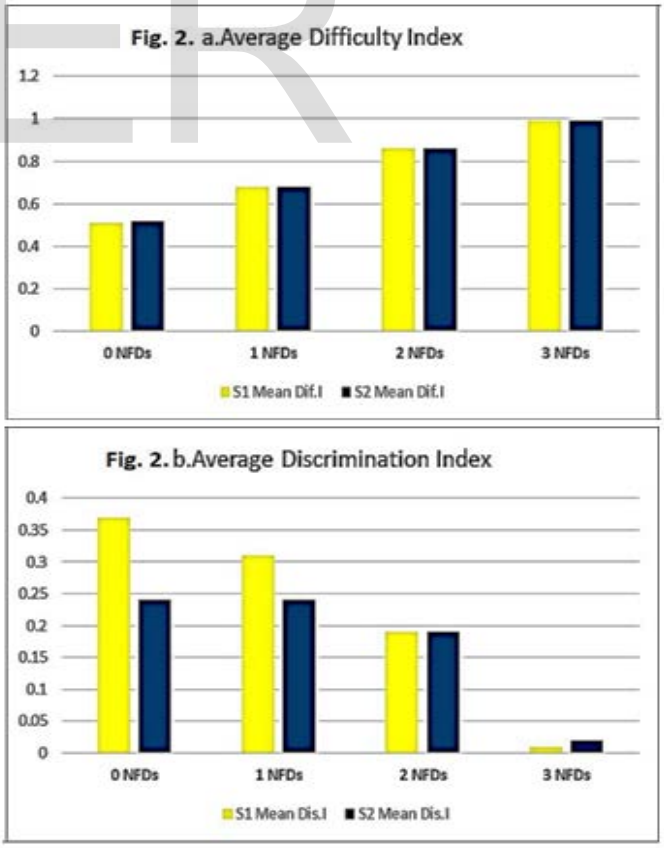

Fig. 2 (a, b): Distribution of difficulty and discrimination indices per number of NFD in both semesters

4 DISCUSSION

4.1 This study is among few large studies of item analysis of MCQs in the Middle East. It is the largest in Egypt and Arabic area. We evaluated 1855 MCQs. By thorough research and literature review, authors found a plethora of articles touching this topic of psychometric analysis of MCQs. However, most of these publications were either presenting a small number of students, small number of questions or using single course as an example for their analysis.

4.2 DIFFICULTY INDEX

The mean difficulty index (Diff I) in our study was 0.71, which means that most of the questions were closer to the easy aspect. The nearest three studies in accordance with our results are Mehta G. (2014) [17], Namdeo SK. (2016) [21] and D'Sa, Juliana (2017) [28]. They recorded average difficulty indices of 0.63, 0.66 and 0.68 respectively. In the current research, there were 46% of the questions in the easy category. However, the previously mentioned three studies reported 32%, 32% and 50% respectively in the easy category. In the current research, the difficult questions represented 11%, whereas, the same previously mentioned studies reported 3%, 8% and 0% respectively in the difficult category. From this comparison, although our results showed higher value of Diff I (easier exam), yet we reported higher percentage of difficult questions which can be of benefit for the higher-level students.

Some studies as Patil VC. (2015). [18], Menon AR. (2017). [23], Gajjar S. (2014) [29] and Suryadevara (2018) [30] reported lower average difficulty index (0.49, 0.45, 0.39, and 0.45 respectively) and higher percentage of difficult questions (22%, 25%, 32% and 18% respectively). Despite these studies reflected a more reasonable level of difficulty index, yet their results were difficult to be extrapolated because they were done on a small number of both students and MCQs.

There are two studies analyzing good number of MCQs. The first, Sayyah, M., et al. (2012) [16] studied a sample of 37 nursing students all courses analyzing 1793 MCQs. The second, Kheyami D. et al. (2018) [26] evaluated 38 students 4 times a year with 50 questions each time, yielding a total number of 800 analyzed MCQs. The two studies reported lower average difficulty index (0.54 and 0.52 respectively) compared to 0.71 in the current study. The percentage of difficult questions were higher (18% and 21% respectively) compared to the 11% in the current study. The higher percentage of easy questions in the current study can be explained on the basis that all students in our study were in the first year, which mandated the use of more level one questions in addition this was the first large scale trial of MCQ testing, so the maturity of the questions and the test makers might still need more effort. These two studies, [16], [26], were done at different student levels, beyond the first year, and their utilization of MCQs was practiced for many years.

4.3 DISCRIMINATION INDEX

The mean discrimination index (Disc I) in the current study was 0.27 with only 15% of the questions having poor discriminatory value. Patil VC. (2015) [18], Shoo, D. (2017) [22] and Gajjar S. (2014) [29] evaluated 100, 60 and 50 MCQs respectively. They reported discrimination index close to the current study findings; 0.19, 0.28 and 0.14 respectively. On the other hand, they reported a higher percentage of questions with poor discrimination 23%, 35% and 52% respectively. Sahoo, D. (2017) [22] , Namdeo SK. (2016) [21], Mehta G. (2014) [17] , Ingale, A. (2017). [24] and Kolte V. (2015) [19] reported higher discrimination index than the current study (0.28, 0.33, 0.33, 0.40 and 0.33 respectively). They also reported very high percentage of poorly discriminating MCQs (35%, 32%, 30%, 27% and 24% respectively). Those studies were based on a small number of questions and this may explain these findings. Also, the higher number of poor MCQs discrimination means less discriminatory power of their tests.

The two largest studies; Sayyah, M., et al. (2012) [16] and Kheyami D. et al. (2018) [26], reported average discrimination index of 0.36 and 0.27 respectively, compared to 0.27 in the current study. The percentage of poor discriminatory questions were higher (25% and 32% respectively) compared to the 15% in the current study. The higher percentage of good discriminatory questions in the current study can be explained in association with the previously mentioned difficulty index. Our current study, despite having a lower percentage of difficult questions yet they discriminated effectively between high scoring students and low scoring ones.

Four studies [20], [23], [25], [28] reported similar discrimination index and less percentage of MCQs with poor discrimination. These studies, despite being based on low number of MCQs, had reported better results than the current study as regard more items with good discrimination and less number of poorly discriminating MCQs.

4.4 NON-FUNCTIONING DISTRACTORS

In the current study, the percentage of MCQs, which had non-functioning distractors (NFDs), were 41% of the total items. This value was the highest if compared to all the reported publications [16], [20], [22] ,[30]. Only, Namdeo SK. (2016) [21] reported NFDs in 53% of the MCQs. This may be due to lower quality of questions with some ambiguous distractors. Most of the studies were done on small number of questions where the pre exam validation and quality control steps of MCQs writing were more feasible. Sayyah, M., et al. (2012) [16] and Kheyami D. et al. (2018) [26] reported NFDs percentage as 34% and 17% respectively. These two studies were done at more mature and well-established systems with good quality control steps and might be due to inclusion of senior student levels. Gajjar S. (2014) [29] reported the lowest non-functioning distractors.

The distribution of NFDs within the MCQs was evaluated. Researches with small MCQ numbers had reasonable good percentage of questions in the 0 NFDs category. Only 3 studies

(Namdeo SK. (2016) [21], Ingale A.(2017) [24] , Mahjabeen,W. [2018) (25] reported less percentage in the items with 0 NFDs. This is expected as in these smaller studies, it is easier to revise and check the quality of such small set of questions.

The current study reported the presence of NFDs as following; 0 NFDs in 30%, 1 NFDs n 40 %, 2 NFDs in 23% and those with >= 3 NFDs in 7%. Compared to the category of large studies, there were two studies. Sayyah, M. (2012) [16] reported the presence of NFDs as following; 0 NFDs in 19%, 1 NFDs in 36 %, 2 NFDs in 28% and those with >= 3 NFDs in 17%. Kheyami D. (2018) [26] reported the presence of NFDs as following; 0 NFDs in 48%, 1 NFDs in 36 %, 2 NFDs in 11% and those with >= 3 NFDs in 5%.as regard the NFDs we had better results than Sayyah, M. (2012) [16], but worse results than Kheyami D. (2018) [26]. The first study (Sayyah M.) was conducted at the same period for all the courses similar to the conditions in our current study while the second study (Kheyami D.) studied questions in four subsequent years. This can explain the improved percentage of good items as a process of system maturity and staff experience.

Distractor efficiency (DE%) in the current study was 56.5%. It was lower than nearly all the compared studies. Again, this reflects the quality of distractors in the current study.

4.5 RELATION BETWEEN THE NUMBER OF OPTIONS AND NFDS, AND THE EFFECT ON PSYCHOMETRIC PARAMETERS

In the current study, as seen from figure 1,and 2, the number of NFDs increased with increasing number of options. The best was 3 option MCQs. This was in accordance with some published articles (Tarrant M (2009) [12], Patil VC. (2015) [18], Mahjabeen,W. (2018) [25] , Patel RM (2017) [31] , Kheyami D. et al. (2018) [26] , Mukherjee P (2015) [32] , Vegada B (2016) [33] and Kenneth Royal (2018) [34] . In addition, the average difficulty index was proportionally increased with increasing NFDs while the discrimination index was proportionally decreasing with increased NFDs. This is can be simple explained by the fact that increasing options will inevitably lead to the inclusion of weaker, odd, non-homogenous or non-plausible options. They will be omitted easily and not selected. Hence, they will increase the NFDs and decrease the distractor efficiency. This will make the MCQs easier (high difficulty index) and less discriminatory (low discrimination index). Saxena S. (2016) [27], Gajjar S. (2014) [29], Kheyami D. et al. (2018) [26] and Mukherjee P (2015) [32] concluded similar results and stated that the more NFDs in an item increases DIF I (makes item easy) and reduces DE, conversely item with more functioning distractors decreases DIF I (makes item difficult) and increases DE.

5 CONCLUSION

This study is among few large studies of item analysis of MCQs in the Middle East. It is the largest in Egypt and Arabic area. The study included 1855 MCQs.
The percentage of non-functioning distractors (NFDs), were 41% of the total items. The highest percentage was 1 NFD in 40 % of items. The number of NFDs increased with increasing number of options. The lesser the number of NFDs, the better the difficulty and discrimination. The best values of NFDs were found in the three options MCQs. The three option MCQs had the best number of good distractors with 0 NFDs, hence they have the reasonable difficulty and best discrimination.

6 LIMITATIONS

In the present study, item analysis and post exam validation were done. Either some items were eliminated or keys were corrected. All the feedback were given to the Faculty members. Despite, many workshops were done prior to the exam on item writing guidelines and post exam validation reports; however, it is evident that the quality of the MCQs and option writing are still need improvement. The program academic committee advised the Faculty to apply the full guidelines. However, some are still in favor of using 4 or 5 options MCQs and this affected the quality by including non -suitable and non-plausible options.

7 ACKNOWLEDGMENT

We offer my unending thanks to faculty members sharing in year one teaching in the CBMBP in addition we also thank Mr. Ahmed Al Fayoumy the technician of the question bank.

8 FINANCIAL SUPPORT AND SPONSORSHIP

There was no support.

9 CONFLICT OF INTEREST

There are no conflicts of interest.

10 REFERENCES

[1] Schuwirth LWT, Van Der Vleuten CPM. Different Written Assessment Methods: What Can Be Said About Their Strengths And Weaknesses? Med Educ. 38(9):974-9, 2004

[2] Kuechler WL, Simkin MG. How Well Do Multiple Choice Tests Evaluate Student Understanding In Computer Programming Classes? J Inform Syst Educ. 14(4):389-400, 2003

[3] Lau PNK, Lau SH, Hong KS, Usop H. Guessing, Partial Knowledge, and Misconceptions in Multiple-Choice Tests. J Educ Technol Soc. 14(4):99, 2011

[4] Haladyna TM, Downing SM, Rodriguez MC. A review of multiple-choice item-writing guidelines for classroom assessment. Applied measurement in education. 15(3):309-34, 2002

[5] Rasiah Si-Mui Sim, Isaiah Raja. Relationship Between Item Difficulty And Discrimination Indices In True/False-Type Multiple Choice Questions Of A Para-Clinical Multidisciplinary Paper. Annals AMS. 35(2):67-71, 2006

[6] Boopathiraj C, Chellamani K. Analysis of Test Items on Difficulty Level and Discrimination Index in The Test For Research In Education. Int J Soc Sci Interdiscip Res. 2(2):189-93, 2013

[7] Mitra N, Nagaraja H, Ponnudurai G, Judson J. The Levels of Difficulty and Discrimination Indices in Type A Multiple Choice Questions Of Pre-Clinical Semester 1 Multidisciplinary Summative Tests. Int e-J Sci, Med & Educ. 3:2-7, 2009

[8] Khafagy, Ghada; ahmed, Marwa; saad, Nagwa. Stepping Up of MCQs' Quality Through A Multi-Stage Reviewing Process. Educ Prim Care, 27(4):299-303, 2016

[9] Matlock-Hetzel S. Basic Concept. In Item And Test Analysis. Presented at annual meeting of the Southwest Educational Research Association, Austin Tx, January; 1997.

[10] Singh T, Gupta P, Singh D. Principles of Medical Education. 4th ed ed: Jaypee Brothers Medical Publishers; p 109, 2013.

[11] Pande SS, Pande SR, Parate VR, Nikam AP, Agrekar SH. Correlation between Difficulty and Discrimination Indices of Mcqs in Formative Exam in Physiology. SE Asian J Med Educ. 7(1):45-50, 2013

[12] Tarrant M, Ware J, Mohammed AM. An Assessment of Functioning and Non-Functioning Distractors in Multiple-Choice Questions: A Descriptive Analysis. BMC med educ. 9(1):40, 2009

[13] Hingorjo MR, Jaleel F. Analysis of One-Best MCQs: The Difficulty Index, Discrimination Index and Distractor Efficiency. JPMA. 62(2):142, 2012

[14] Ananthakrishnan N. Item Analysis-Validation And Banking Of MCQs. N Ananthkrishnan, KR Sethuraman & S Kumar Medical Education Principles and Practice Pondichery: JIPMER. 131-7, 2000

[15] Singh JP, Kariwal P., Gupta SB., Shrotriya VP. Improving Multiple Choice Questions (MCQs) Through Item Analysis: An Assessment of The Assessment Tool. IJSAR. 1(2):53-7, 2014

[16] Sayyah M, Vakili Z, Masoudi Alavi N, Bigdeli M, Soleymani A, Assarian M, et al. An Item Analysis of Written Multiple-Choice Questions: Kashan University of Medical Sciences. Nurs Midwifery Stud. 1(2):83-7, 2012

[17] Mehta G, Mokhasi V. Item Analysis of Multiple Choice Questions-An Assessment of The Assessment Tool. Int J Health Sci Res. 4(7):197-202, 2014

[18] Patil V, HV P. Item Analysis of Medicine Multiple Choice Questions (MCQs) For Under Graduate (3rd Year MBBS) Students. Res J Pharm Biol Chem Sci. 6(3):1242-51, 2015

[19] Kolte V. Item Analysis of Multiple Choice Questions in Physiology Examination. Indian J Basic Appl Med Res. 4(4):320-6, 2015

[20] Kaur M, Singla S, Mahajan R. Item Analysis Of In Use Multiple Choice Questions In Pharmacology. Int J Appl Basic Med Res. 6(3):170–3, 2016

[21] Namdeo SK, Sahoo B. Item Analysis of Multiple Choice Questions from An Assessment of Medical Students In Bhubaneswar, India. Int J Res Med Sci. 4(5):1716-9, 2016

[22] Sahoo DP, Singh R. Item and Distracter Analysis of Multiple Choice Questions (MCQs) From A Preliminary Examination of Undergraduate Medical Students. Int J Res Med Sci. 5(12):5351-5, 2017

[23] Menon AR, Kannambra PN. Item Analysis To Identify Quality Multiple Choice Questions. Nat J of Lab Med. 2017;6(2):MO07-10.

[24] Ingale AS, Giri PA, Doibale MK. Study on Item and Test Analysis of Multiple Choice Questions Amongst Undergraduate Medical Students. Int J Comm Med Pub Health. 4(5):1562-5, 2017

[25] 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. 13(4):310-5, 2018

[26] Kheyami D, Jaradat A, Al-Shibani T, Ali FA. Item Analysis of Multiple Choice Questions At The Department Of Paediatrics, Arabian Gulf University, Manama, Bahrain. S Q U Med J. 18(1):e68, 2018

[27] Saxena S., Srivastava PC, Mallick AK, Joshi HS, and Das B. Item Analysis: An Unmatched Tool For Validating MCQs In Medical Education. Indian J Basic Appl Med Res. 5(4):263-69, 2016

[28] D'Sa JL, Visbal-Dionaldo ML. Analysis of Multiple Choice Questions: Item Difficulty, Discrimination Index and Distractor Efficiency. Int J Nurs Educ. 9(3), 2017

[29] Gajjar S, Sharma R, Kumar P, Rana M. Item and Test Analysis to Identify Quality Multiple Choice Questions (MCQs) From An Assessment Of Medical Students Of Ahmedabad, Gujarat. Indian J Comm. Med. 39(1):17, 2014

[30] Suryadevara VK, Bano Z. Item Analysis To Identify Quality Multiple Choice Questions/Items In An Assessment In Pharmacology Of II MBBS Students In Guntur Medical College Of Andhra Pradesh, India; 2018.

[31] Patel RM. Use of Item Analysis to Improve Quality Of Multiple Choice Questions In II MBBS. J Educ Technol Health Sci. 4(1):22-9, 2017

[32] Mukherjee P, Lahiri SK. Analysis Of Multiple Choice Questions (MCQs): Item And Test Statistics From An Assessment In A Medical College Of Kolkata, West Bengal. IOSR J Dent Med Sci. 1:47-52, 2015

[33] Vegada B, Shukla A, Khilnani A, Charan J, Desai C. Comparison Between Three Option, Four Option And Five Option Multiple Choice Question Tests For Quality Parameters: A Randomized Study. Indian J Pharm. 48(5):571, 2016

[34] Royal K, Dorman D. Comparing Item Performance on Three-Versus Four-Option Multiple Choice Questions in A Veterinary Toxicology Course. Vet Sci. 5(2):55, 2018

IJSER