

Nonfunctional distractor analysis: An indicator for quality of Multiple choice questions

Madiha Sajjad¹, Samina Iltaf², Rehan Ahmed Khan³

ABSTRACT

Objectives: To analyze the low to medium distractor efficiency items in a multiple-choice question (MCQ) paper for item writing flaws.

Methods: This qualitative study was conducted at Islamic International Medical College Rawalpindi, in October 2019. Archived item- analysis report from a midyear medium stakes MCQ paper of 2nd year MBBS class, was analyzed to determine the non-functional distractors (NFDs) and distractor efficiency (DE) of items, in a total of 181 MCQs. DE was categorized as low (3-4 NFDs), medium (1-2 NFDs) and high (0 NFD). Subsequently, qualitative document analysis of the MCQ paper whose item analysis report was assessed was conducted to investigate the item flaws in the low to medium DE items. The flaws identified were coded and grouped as, within option flaws, alignment flaws between options and stem/ lead-in and other flaws.

Results: Distractor efficiency was high in 69 items (38%), moderate in 75 items (42%) and low in 37 items (20%). The item-writing flaws identified in low to moderate DE items within distractors included, non-homogenous length (1.8%), non-homogenous content (8%) and repeat in distractor (1.7%). Alignment flaws between distractors and stem/ lead-in identified were linguistic cues (10%), logic cues (12.5%) and irrelevant distractors (16%). Flaws unrelated to distractors were low cognitive level items (40%) and unnecessarily complicated stems (11.6%).

Conclusions: Analyzing the low to medium DE items for item writing flaws, provides valuable information about item writing errors which negatively impact the distractor efficiency.

KEYWORDS: Non-functional distractor, Distractor efficiency, Multiple choice question, Item writing flaws.

doi: <https://doi.org/10.12669/pjms.36.5.2439>

How to cite this:

Sajjad M, Iltaf S, Khan RA. Nonfunctional distractor analysis: An indicator for quality of Multiple choice questions. *Pak J Med Sci.* 2020;36(5):982-986. doi: <https://doi.org/10.12669/pjms.36.5.2439>

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

INTRODUCTION

Post hoc item analysis is a commonly used tool to assess the quality of Multiple-choice questions (MCQ) based examinations in undergraduate medical education. It provides useful information about the reliability and validity of test items. The parameters commonly assessed in item analysis are; the discrimination index (DI), difficulty index and distractor efficiency (DE). Functional or efficient distractors are those, chosen by more than 5% of examinees whereas the distractors chosen by less than 5% examinees are known as non-functional distractors.¹ For distractors to be effective they should all be plausible and if possible, none should be incorrect.² The number of non-functional distractors NFDs in an MCQ item

1. Dr. Madiha Sajjad, FCPS.
Department of Pathology,
2. Dr. Samina Iltaf, MPhil.
Department of Pathology,
3. Dr. Rehan Ahmed Khan, FCPS, FRCS, JM-HPE, MSc HPE, PhD.
Department of Surgery,
- 1-3: Riphah International University,
Rawalpindi, Pakistan.

Correspondence:

Prof. Dr. Rehan Ahmed Khan,
Surgery Department,
Riphah International University,
Rawalpindi, Pakistan.
Email: surgeonrehan@gmail.com

- * Received for Publication: February 29, 2020
- * Revision Received: May 19, 2020
- * Revision Accepted: May 22, 2020

determines the distractor efficiency of that item. Designing plausible distractors and reducing the number of non-functional distractors (NFDs) improves the quality of the test.

In an MCQ item, the context or background is referred to as the 'stem', followed by the question known as the 'lead-in' and a number of option choices. High-quality MCQs require a well written unambiguous stem, clear lead-in and rational choice of options. In a one best type of MCQ, one of the options is the correct response known as the 'key' while others are described as 'distractors'.³ Item writing flaws in MCQs which are not related to item construct occur when there is a breach in following the standard item-writing guidelines with reference to language and grammatical structure, style of writing the stem and option choices.⁴ Various types of item writing flaws are found in literature, for example; Long correct answer, logic cues, grammar cues, 'except' or 'not' in the lead-in, inconsistent language in options etc.⁵ Some types of Item flaws cue the student to the correct answer, assisting students who are 'test wise'. 'Test-wiseness' refers to students' ability to recognize the answer in MCQs without employing their content-related reasoning skills or knowledge. Other types of flaws may mislead the students towards selecting wrong options.⁴ This may over or under-estimate student performance, thus introducing a source of error that negatively effects the validity of student scores.^{4,6}

In depth study of items showing low distractor efficiency can help test-developers and instructors understand test-wiseness of students in relation to item flaws. Distractor efficiency also has an indirect effect on the item difficulty as well as its discriminatory ability.⁷ Removing the non-functional distractors from MCQ items in some cases, restores the distractor efficiency of item to optimal level. In other cases, identifying distractor related flaws and correcting them can improve distractor efficiency and positively impact the item difficulty and discrimination indices.⁸

Not much work has been done on the qualitative aspect of individual items with low distractor efficiency. The objective of this study was to analyze the item writing flaws in low to medium distractor efficiency items in a multiple-choice question (MCQ) paper in order to gain insight into the structural flaws in items which negatively impact the distractor efficiency,

overall exam quality and student performance. This study will help the test developers be aware of item flaws and address them in a more logical and systematic manner.

METHODS

This qualitative study was conducted at Islamic International Medical College, Riphah International University in October 2019. Ethical approval was taken from Institutional review committee of Islamic International Medical college, Ref #Riphah/IIMC/IRC/20/005.

Archived item- analysis data report from a midyear medium stakes MCQ paper of 2nd year MBBS class was analyzed to assess the distractor efficiency of items. There were 181 single best response type MCQs having a reliability coefficient of 0.88. One hundred and six (106) items had five option choices and 75 items had four option choices.

Nonfunctional distractors (NFD) were identified as the distractors chosen by less than 5% examinees. Distractor efficiency (DE) was defined on the basis of the number of NFDs in an item and ranged from 0-100%. Distractor efficiency of items was graded as low (having 3-4 NFDs), medium (having 1-2 NFDs) and high (having 0 NFD).

Subsequently, qualitative document analysis of the MCQ paper whose item analysis report was assessed, was carried out independently by two reviewers, who evaluated each low to moderate DE item, for item writing flaws. The reviewers were experienced in MCQ test item development and were also trained in MCQ writing. Item flaws were investigated with reference to item writing guidelines proposed by Haladyna et al. and followed in literature.⁹⁻¹¹ After consensus, the flaws identified were coded and grouped as:

- Flaws within options
- Alignment flaws between options and stem/ lead-in
- Other flaws, unrelated to options or their alignment with stem/ lead-in

RESULTS

A total of 649 distractors were identified in 181 MCQs out of which 205 were nonfunctional distractors (31.6%). Out of these 181 MCQ items, 112 items were low to moderate distractor efficiency items. Distractor efficiency of items was determined as given in Table-I.

Table-I: Distractor efficiency of items.

No. of Non-functional distractors (NFD's)		Distractor efficiency	No. of items
4 option items	5 option items		
0	0	High (100%)	69 items (38%)
1-2	1-2	Moderate (50-75%)	75 items (42%)
3	3-4	Low (< 50%)	37 items (20%)

Out of the 112 low to moderator distractor efficient items, 62 items (34.2% of all MCQ items) had one or more item writing flaws in the option choices or their alignment with the stem and lead-in. Of the remaining 50 items; flaws unrelated to the options or their alignment with stem/ lead-in were found in 27 items, these were low cognition level items or complicated stems; whereas no item writing flaw was identified in 23 items. The item flaws identified are given in Table-II.

The most frequently identified flaws were 'low cognitive level items' (40%), 'irrelevant/ limited possible options or distractors' (16%), followed by 'logic cues' and (12.5%) and, unnecessarily complicated stems (11.6%).

DISCUSSION

We aimed to identify different types of qualitative item flaws in MCQ items having low to moderate distractor efficiency. Distractor efficiency was low to moderate in 62% items in our study. Various studies show comparable results with low to moderate distractor efficient items in the range of 31% to 75% in various local studies,^{1,12,13} and 50-86% in international studies.^{6,14}

The frequency of flawed items in the 112 low to moderate distractor efficient items in our study was 89/112, (79%) which was 49% of the total MCQ items. In a study by Pais et al., 55.8% items had at least one item writing flaw. This high frequency was similar to other studies, where

around half of the items contained item writing flaws.^{5,6,11,15,16} Flaws in items included limited plausible distractors, clues, unfocused stems, errors in writing option choices or those related to cognitive level chosen etc. These flaws may cue the students and cause the distractors to be chosen by students based on their 'guessing' skills rather than content-specific cognitive skills.^{11,17}

Lower-order cognition items were a frequently identified flaw (40%). Ideally constructed MCQs should be written at a level of difficulty appropriate to level of the students and the focus of assessment should not be students' knowledge of inconsequential or trivial facts.¹⁸ The low cognition item stems based on recall are sometimes too easy for the level/ grade of students, causing students to not consider any distractor as an option. In a study by Testa et al, items categorized at the 'Application level' were more distractor efficient when compared to items labelled at 'Knowledge' and 'Comprehension' levels.¹⁹ In multiple studies low cognition items were in the range of 40-60%, as was the case in our study.^{11,20}

The predominant item flaw identified following the 'low cognition items' was 'limited number of possible logical or plausible distractors' (16%) which was in the category Alignment flaws between the distractors and the stem or lead-in'. In a study by Salih et al., implausible distractors were cited as a frequent flaw (25%).²⁰ Pham et al. observed that students who can 'rule out' options based on their ability to assess their plausibility

Table-II: Frequency of item writing flaws in low and moderate distractor efficiency items.

Within options flaws (11.5%)		Alignment flaws between options and item stem/ lead in (38.5%)		Other flaws	
Distractor non-homogenous in length	1.8% (3)	Linguistic cues	10% (11)	Low cognitive level items	40% (45)
Distractor non-homogenous in content	8% (9)	Logic cues	12.5% (14)	Unnecessarily complicated/ unfocused stem	11.6% (13)
Distractor repeated in same item	1.7% (2)	Limited possible or irrelevant distractors	16% (18)	No item writing flaw	20.5% (23)

can narrow down their guessing to fewer options than the five which are frequently employed in MCQs.⁴ Test developers are generally asked to provide four or five option choices and it may be difficult for them to furnish quality distractors every time, as some questions inherently have less plausible options. According to studies by and Tarrant et al. and Haladyna et al., the quality of distractors rather than their number is the crucial requisite and suggested the minimum required number of options in an item should be considered in the context of the ability to develop plausible distractors even reducing the number of options to just three.^{6,9}

Another frequently identified flaw in our study was 'logic cues' due to inadvertent use of specific words or themes in distractors and stem or lead-in directing towards the correct answer (12.5%). In some other studies these were less frequent (4-6%).^{5,11}

Ambiguous/ confusing stem or lead-in was found to be 11.6% in our study. In various studies the frequency of this flaw ranged from 19-50%.^{5,11,20} Vague or ambiguous terms effect the ability of students to answer a question correctly.¹⁸ One of the reasons for these flaws in our study may be that English is not the primary language in the study setting, although it is the medium of instruction (EMI).⁶ The impact of language on the psychometric properties, due to item quality as well as the examinees' comprehension needs to be further studied.

The least frequent flaws in our study were in the category, 'within options flaws' (11.5%). This was in accordance with the study by Salih et al.²⁰ However other studies show a higher frequency of these flaws.¹¹ The reason for a lower frequency of 'within option flaws' in our study may be, because of a system of pre-hoc analysis for medium and high stakes examination items at our institution. This may lead to reduction in such obvious flaws in item writing as, 'all of the above' or 'none of the above' options etc.²⁰ Also, having a faculty development program in place to train faculty in standard item writing guidelines also helps improve the quality of MCQs.¹¹

No item writing flaw was identified in 20% items having low to moderate distractor efficiency. One of the possible reasons may be that the item psychometric statistics from only one administration of the test was studied, and the

options identified as non-functioning distractors in this exam may have been subject to sampling bias and may perform quite differently in other samples.⁶

Limitation of the study: A limitation of the study is, that a single paper was analyzed for item flaws based on its item analysis index of low DE.

CONCLUSION

Analyzing the low to medium DE items for item writing flaws, provides valuable information about item writing errors which negatively impact the distractor efficiency. Correcting the errors can improve distractor efficiency and overall exam quality.

Grant Support & Financial Disclosures: None.

REFERENCES

1. Iqbal MZ, Khan RA, Razaq N. Assessment of non-functional distractors in multiple choice questions: A descriptive analysis. *Pak J Physiol.* 2016;12(2):47-49.
2. Burud I, Nagandla K, Agarwal P. Impact of distractors in item analysis of multiple choice questions. *Int J Res Med Sci.* 2019;7(4):1136-1139. doi: 10.18203/2320-60
3. Rahma A, Shamad M, Idris MEA, Elfaki O, Elfakey W, Salih KMA. Comparison in the quality of distractors in three and four options type of multiple choice questions. *Adv Med Educ Pract.* 2017;8:287-291. doi: 10.2147/amep.s128318
4. Pham H, Besanko J, Devitt P. Examining the impact of specific types of item-writing flaws on student performance and psychometric properties of the multiple choice question. *MedEdPublish.* 2018;7(4):1-16. doi: 10.15694/mep.2018.0000225.1
5. Kowash M, Hussein I, Halabi M Al. Evaluating the quality of multiple choice question in paediatric dentistry postgraduate examinations. *SQU Med J.* 2019;19(2):135-141. doi: 10.18295/squmj.2019.19.02.009
6. Tarrant M, Ware J, Mohammed AM. An assessment of functioning and non-functioning distractors in multiple-choice questions: A descriptive analysis. *BMC Med Educ.* 2009;9(40). doi: 10.1186/1472-6920-9-40
7. Abdulghani H, Ahmad F, Aldrees A, Khalil M, Ponnamparuma G. The relationship between non-functioning distractors and item difficulty of multiple choice questions: A descriptive analysis. *J Heal Spec.* 2014;2(4):148-151. doi: 10.4103/1658-600X.142784
8. Ali SH, Ruit KG. The Impact of item flaws, testing at low cognitive level, and low distractor functioning on multiple-choice question quality. *Perspect Med Educ.* 2015;4(5):244-251. doi: 10.1007/s40037-015-0212-x
9. Haladyna TM, Downing SM, Rodriguez MC. A review of multiple-choice item-writing guidelines for classroom assessment. *Appl Meas Educ.* 2002;15(3):309-334. doi: 10.1207/S15324818AME1503_5
10. Haladyna T, Rodriguez MC. *Developing and validating test items.* 1st ed. New York; 2013:454. doi: 10.4324/9780203850381 p

11. Tariq S, Tariq S, Maqsood S, Jawed S, Baig M. Evaluation of cognitive levels and item writing flaws in medical pharmacology internal assessment examinations. *Pak J Med Sci.* 2017;33(4):866-870. doi: 10.12669/pjms.334.12887
12. Hingorjo MR, Jaleel F. Analysis of one-best MCQs: The difficulty index, discrimination index and distractor efficiency. *J Pak Med Assoc.* 2012;62(2):142-147.
13. 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.* 2017;5(12):5351-5355. doi: 10.18203/2320-6012.ijrms20175453
14. D'Sa JL, Dionaldo MLV-. Analysis of multiple choice questions: Item difficulty, discrimination index and distractor efficiency. *Int J Nurs Educ.* 2017;9:109-114. doi: 10.5958/0974-9357.2017.00079.4
15. Pais J, Silva A, Guimaraes B, Povo A, Coelho E, Pereira FS, et al. Do item - writing flaws reduce examinations psychometric quality? *BMC Res Notes. BioMed Central;* 2016;9:399. doi: 10.1186/s13104-016-2202-4
16. Tarrant M, Knierim A, Hayes SK, Ware J. The frequency of item writing flaws in multiple-choice questions used in high stakes nursing assessments. *Nurse Educ Today.* 2006;26(8):662-671.
17. Royal KD, Hedgpeth M-W, Posner LP. A simple methodology for discerning item construction flaws in health professions examinations. *Heal Prof Educ. Elsevier B.V.* 2019;5(1):82-89. doi: 10.1016/j.hpe.2018.03.001
18. Al-Rukban MO. Guidelines for the construction of multiple choice questions tests. *J Family Community Med.* 2006;13(3):125-133.
19. 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. *Front Psychol.* 2018;9:1-12. doi: 10.3389/fpsyg.2018.01585
20. Salih KMA, Elgusadi MA, Elfaki OA. High thinking multiple choice questions for Pediatrics II course at King Khalid University, Kingdom of Saudi Arabia. *JSMSB.* 2017;1(1):3-6.

Authors' Contribution:

RAK: Conceived, designed reviewed and did final approval.

MS and SI: Did data collection, manuscript writing, editing and review.

All authors are responsible and accountable for the accuracy and integrity of the work.