

Original article:

Analysis and remedy of the item writing flaws rectified at prevalidation of multiple choice questions drafted for assessment of MBBS students

Sood RS¹, Bendre MB², Sood A³

¹Associate Professor, Physiology, Dr. DY Patil Medical College, Hospital & Research Centre, Pune, Maharashtra

²Professor, Surgery, Dr. DY Patil Medical College, Hospital & Research Centre, Pune, Maharashtra

³ Medical Practitioner, Trust Clinic, Subhash Nagar, Vishrantwadi, Pune, Maharashtra

Corresponding author: Dr Sood RS

Abstract:

Introduction: An examination based on Multiple Choice Questions (MCQs) serves as an excellent tool of assessment. This modality eliminates subjectivity and makes scoring of answer sheets time efficient. It is, thus, not surprising that this modality is used frequently. MCQ may lose its value altogether due to presence of Item Writing Flaws (IWFs) and must be designed well for examination to remain valuable as a truthful assessment tool. Therefore, it is a routine practice to subject MCQs to prevalidation by experienced faculty, before including these in the question paper. This study looked at the IWFs detected at the stage of prevalidation, with a view to recommend measures to prevent the same.

Methods: One thousand draft MCQs were prevalidated. The IWFs were classified, tabulated and critically analyzed wherever required. The precise cause of the IWFs was arrived at and suitable rectificatory measures listed.

Observations: IWFs were discovered in 45% of the MCQs. Flawed MCQs mostly (84%) had a solitary IWF with only 15% had two each. More than two flaws in an MCQ were rare.

Results: Most of the IWFs were traceable to carelessness with the usage of language. Only 7% were traceable to suboptimal knowledge of content.

Conclusion: The single most effective measure to prevent IWFs is sensitization of the faculty to the issue. Due importance must be given to the optimal use of language. Content knowledge of the faculty was not found to have been a major issue.

Key words: Item Writing Flaws, Multiple Choice Questions, Prevalidation

Introduction:

In ancient Greece, Socrates tested his students through conversations, the 'Socratic method'. Frederick J. Kelly, University of Kansas, was the first to use multiple choice questions (MCQs) as part of a large scale assessment in 1914. A strong correlation has been reported between the scores of MCQs and essay questions.^[1,2] Well-constructed MCQs are considered a satisfactory replacement and even superior to essay questions in testing the higher cognitive skills of undergraduate medical students

and high stakes exit examinations.^[3-5] Use of well-constructed MCQs has been reported the most accurate measure of student achievement compared with other forms of evaluations including written assignments.^[6]

Advantages of this method include ease of administration, even to a large group of students, avoidance of subjectivity and prompt optical reader aided scoring of the answers. As the factors irrelevant to the assessed material, such as handwriting and

clarity of presentation, do not come into play, grading is based purely on the knowledge of the topic.

Administration of examinations based on MCQs, thus, remains an efficient method of assessment at various entrance examinations as well as the formative and summative assessments during MBBS course.^[1,3,4,6] However, to remain effective, it is of paramount importance that a high standard of quality of MCQs be maintained. Those with item writing flaws (IWFs) are unlikely to be good assessment tools.^[7,8] Flawed items have been reported 0-15 percentage points more difficult than standard items measuring the same construct.^[9] Thus high-achieving students have been reported more likely than borderline students to be penalized by such items.^[10] IWFs have been reported even in the MCQs of continuing medical education (CME) activities published by major medical journals.^[11-13] Prevalidation has been recommended to reduce the occurrence of IWFs and to improve the quality of test questions.^[14] Guidelines are available for sound construction of MCQs and must be exploited.^[5,15-18] Sensitization of

faculty to the issue of occurrence of various types of IWFs and possible means to avoid those go a long way in maintaining a high standard & thus utility of MCQs.^[19-21]

Aims & Objectives:

This study aimed to analyze the occurrence of IWFs, classify, find the causes and recommend means to avoid the same while constructing MCQs.

Material & Methods:

One thousand MCQs drafted by faculty, for formative assessment of I MBBS students in Physiology, were subjected to prevalidation. All the MCQs were supposed to have the question (stem), followed by four choices of answers (options), amongst which there were to be one correct answer (key) and three incorrect answers (distracters). The IWFs noted were classified & tabulated. This study was initiated under the guidance of Medical Education Technology Unit, Dr. DY Patil Medical College, Hospital & Research Centre, Pune, as part requirement of 'Basic Course Workshop in Medical Education Technologies'.

Indian Journal of Basic and Applied Medical Research

Is now with

IC Value 91.48

Observations & Results:

Without any IWFs	549 (55%)
Flawed	451 (45%)
Total	1000 (100%)

MCQs with a solitary IWF	379 (84%)
MCQs with two IWFs	68 (15%)
MCQs with three IWFs	3 (<1%)
MCQs with four IWFs	1 (<1%)
Total	451 (100%)

Stem:	
Not matching the options grammatically	16 (03)
Not clear-cut	61 (12)
Negative statement not highlighted	26 (05)
Options (general):	
Repeated words	168 (32)
Not following an ascending/ descending order for values	10 (02)
Options (key):	
Other than one key	77 (15)
Values not from a standard textbook	05 (<1)
Not well deliberated	05 (<1)
Narrowed down to two options	02 (<1)
Options (distracter):	
Ruled out/ synonymous/ obvious/ ineffective to be useful	19 (04)
Miscellaneous:	
Spelling mistakes	40 (08)
Suboptimal usage of language	94 (18)
Does not test any important issue	05 (<1)
Total	528 (100)

Discussion:

Forty five percent MCQs revealed IWFs. Our results fare better than 46% - 47% reported by Tarrant M, et al^[10, 22], 65% by Downing SM^[9], 68% by Kühne-Eversmann L, et al^[12] and 100% by Stagnaro-Green AS, et al.^[13] Our results are marginally adverse compared to 43% reported by DiSantis DJ.^[11] 15% of the flawed MCQs revealed multiple IWFs.

The single most common (32%) IWF was 'repeated words' in the options. In most (87%) instances these repeated words could be easily incorporated in the stem, though a few (7%) necessitated ingenuity. Thus, very awareness of this type of IWF will keep it from creeping in. This underlines the need for sensitization of the faculty. On about 6% occasions, 'repeated words' could be totally dispensed with.

The second most common IWF (18%) was 'suboptimal usage of language'. Poorly constructed sentences accounted for most (67%) instances of this type of error. This type of IWF showing 'awkward construction' has also been reported by Gutmann A, et al, in the MCQs included in CME of a journal.^[23] The rest were due to grammatical errors (26%) & presence of redundant words (7%). Fifteen percent cases of grammatical errors ended up giving away the key.

15% MCQs had other than one key. There were instances of no key (19%), two keys (61%), three keys (8%) & all four keys (12%). The reasons for this type of IWF included suboptimal knowledge or comprehension of content (91%), carelessness in framing the MCQs (4%), suboptimal use of language (4%) & values taken from non-standard textbooks (1%).

Stem 'not clear-cut' accounted for 12% of the IWFs. Most important cause (75%) of this type of IWF was found to be just poor construction of stem, not

originating from suboptimal content knowledge. Only in 25% cases it was due to suboptimal content knowledge. In two instances, poor construction of the stem even gave away the key.

Spelling mistakes accounted for 8% of the IWFs, occurring more often (60%) in technical terms. The rest (40%) mistakes related to common English & seemed to be attributable to 'autocorrect function of Windows' or typographic error. All these can be avoided by adding these terms to the 'Windows dictionary' and a thorough proof reading. Inability to remove a misspelt word inadvertently added to Windows dictionary remains a problem.

Another 5% of the IWFs were attributed to 'negative statement not highlighted' in the stem. Mere awareness of this type of IWF will prevent it.

Distracter being 'ruled out/ synonymous/ obvious/ ineffective to be useful' contributed 4% IWFs. Most (68%) were due to suboptimal understanding of the content, some (21%) due to just poor deliberation while constructing and a few (11%) due to typographical error. Such IWFs comprising 'Non-functioning Distractors' were reported in 58% MCQs by Hingorjo MR, et al and the author has recommended three functioning distractors should be incorporated in MCQs.^[24]

The IWF 'Stem not matching the options grammatically' accounted for 3% of the IWFs. However, at times the stem did not match many or even all the distracters, turning this IWF innocuous enough to be a problem. This type of IWF showing 'unintentional cues' has also been reported by Gutmann A, et al, in the MCQs included in the CME of a journal.^[23]

Options 'not following an ascending/ descending order for values' accounted for 2% of the IWFs. Giving values in a particular order prevents avoidable

confusion as well as loss of valuable examination time. This type of IWFs will also be prevented by just being aware.

'Not well deliberated' key accounted for less than 1% of the IWFs. These were due to suboptimal content understanding (60%), suboptimal usage of language (20%) & suboptimal deliberation (20%).

MCQ 'Does not test any important issue' also accounted for less than 1% of the IWFs. However this type of IWF assumes significance and must be kept in mind while composing MCQs for specialist CME in print media that need to be based on key message of the CME.^[25]

Conclusion:

Occurrence of IWFs is a reality. If not careful, the composer of MCQs may inadvertently introduce these.^[11-13] It is important to pay attention to not only the content of MCQ but also the correct usage of the

language and proper deliberation to avoid IWFs. Most of the IWFs can be avoided by sensitizing the composer to the issue, paying attention to the language, avoiding carelessness, using values from the standard textbooks, using 'spellcheck' of 'Windows', adding technical terms to 'Windows' dictionary and a thorough proof reading. Prevalidation of MCQs serves to improve the quality. Mere sensitizing the composer to the issue will ward off majority of the IWFs. With that end in mind, faculty must be exposed to faculty development programs to sensitize them to the common IWFs, nuisance thereof and the relative ease with which these can be avoided.^[14]

Acknowledgements:

Medical Education Technology Unit, Dr. DY Patil Medical College, Hospital & Research Centre, Pune.

Reference:

1. Pepple DJ, Young LE, Carroll RG. A comparison of student performance in multiple-choice and long essay questions in the MBBS stage I physiology examination at the University of the West Indies (Mona Campus). *Adv Physiol Educ.* 2010; 34(2): 86-9.
2. Oyebola DD, Adewoye OE, Iyaniwura JO, Alada AR, Fasanmade AA, Raji Y. A comparative study of students' performance in preclinical physiology assessed by multiple choice and short essay questions. *Afr J Med Med Sci.* 2000; 29(3-4):201-5.
3. Khan MU, Aljarallah BM. Evaluation of Modified Essay Questions (MEQ) and Multiple Choice Questions (MCQ) as a tool for Assessing the Cognitive Skills of Undergraduate Medical Students. *Int J Health Sci (Qassim).* 2011; 5(1): 39-43.
4. Palmer EJ, Devitt PG. Assessment of higher order cognitive skills in undergraduate education: modified essay or multiple choice questions? *BMC Med Educ.* 2007; 7:49.
5. Redmond SP, Hartigan-Rogers JA, Cobbett S. High time for a change: psychometric analysis of multiple-choice questions in nursing. *Int J Nurs Educ Scholarsh.* 2012; 9(1): published online.
6. Bontis N, Hardie T, Serenko A. Techniques for assessing skills and knowledge in a business strategy classroom *Int. J. Teaching and Case Studies,* 2009; 2 (2): 162-80.
7. Ware J, Vik T. Quality assurance of item writing: during the introduction of multiple choice questions in medicine for high stakes examinations. *Med Teach.* 2009; 31(3):238-43.

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):244-51.
9. Downing SM. The effects of violating standard item writing principles on tests and students: the consequences of using flawed test items on achievement examinations in medical education. *Adv Health Sci Educ Theory Pract.* 2005; 10(2):133-43.
10. Tarrant M, Ware J. Impact of item-writing flaws in multiple-choice questions on student achievement in high-stakes nursing assessments. *Med Educ.* 2008; 42(2):198-206.
11. DiSantis DJ, Ayoob AR, Williams LE. Journal Club: Prevalence of flawed multiple-choice questions in continuing medical education activities of major radiology journals. *AJR Am J Roentgenol.* 2015; 204(4):698-702.
12. Kühne-Eversmann L, Nussbaum C, Reincke M, Fischer MR. CME activities of medical journals: quality of multiple-choice questions as evaluation tool. Using the example of the German medical journals *Deutsches Arzteblatt*, *Deutsche Medizinische Wochenschrift*, and *Der Internist*. *Med Klin (Munich).* 2007; 102(12):993-1001.
13. Stagnaro-Green AS, Downing SM. Use of flawed multiple-choice items by the *New England Journal of Medicine* for continuing medical education. *Med Teach.* 2006; 28(6):566-8.
14. 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 Pract.* 2006; 6(6):354-63.
15. Mohammed O. Al-Rukban. Guidelines for the construction of multiple choice questions tests. *J Family Community Med.* 2006; 13(3): 125–133.
16. Item Writing Manual by the National Board of Medical Examiners, USA. Available at URL: <http://www.nbme.org/publications/item-writing-manual-download.html>. Accessed on 30 Aug 2015.
17. Caldwell DJ, Pate AN. Effects of question formats on student and item performance. *Am J Pharm Educ.* 2013; 77(4):71.
18. Collins J. Education techniques for lifelong learning: writing multiple-choice questions for continuing medical education activities and self- assessment modules. *Radiographics.* 2006; 26(2):543-51.
19. Abdulghani HM, Ahmad F, Irshad M, Khalil MS, Al-Shaikh GK, Syed S et al. Faculty development programs improve the quality of Multiple Choice Questions items' writing. *Sci Rep.* 2015; 5:9556.
20. Naeem N, van der Vleuten C, Alfaris EA. Faculty development on item writing substantially improves item quality. *Adv Health Sci Educ Theory Pract.* 2012; 17(3): 369-76.
21. Baig M, Ali SK, Ali S, Huda N. Evaluation of Multiple Choice and Short Essay Question items in Basic Medical Sciences. *Pak J Med Sci.* 2014; 30(1):3-6.
22. 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):662-71.
23. Gutmann A, Degirmenci U, Kreil S, Kornhuber J, Weih M. Improvement of CME questions from *Der Nervenarzt*. *Nervenarzt.* 2010; 81(11):1363-7.
24. Hingorjo MR, Jaleel F. Analysis of one-best MCQs: the difficulty index, discrimination index and distractor efficiency. *J Pak Med Assoc.* 2012; 62(2):142-7.

25. Nobbe H, Lösche P, Griebenow R. Structure of text and multiple-choice questions in continuing medical education (CME) in two specialist journals. *Med Klin (Munich)*. 2008;103(1):14-9.