Assessment of distracters in multiple-choice tests

Maryam-Sadat Kaveh Tabatabaee, MSc[1]; Akbar Derakhshan, MD[2]

[1] Faculty Member of Nursing-Midwifery School, Mashad University of Medical Sciences
[2] Director of Education Development Center, Mashad University of Medical Sciences

ABSTRACT

Background *The aim of using distracters in multiple-choice questions (MCQs) is to distract those who don't know the correct answer of the question. The MCQs must be designed in a way that the students who have reached the pre-determined educational objectives, have a better chance in finding the correct answer than those who have not. Therefore, a question has an appropriate efficacy when more low-grade students are attracted to the distracters than high-grade students.*

Purpose *Considering the important role of distracters in differentiating between low and high-grade students, this research studies the distracters in the multiple-choice tests (MCTs) given in Mashad Medical School in the first semester of 2001-2002.*

Method *In this descriptive study, the data of 21 MCTs performed in one semester were collected and recorded by Optical Mark Reader (OMR). From 1284 questions, 642 MCQ questions were sampled randomly, and the numbers of low and high-grade students choosing the distracters were compared and analyzed using SPSS software.*

Results *The efficacy of 45.9% of distracters was "relatively weak" to "extremely weak", with "relatively weak", "weak", "very weak" and "extremely weak" distractors having a prevalence of 11.4%, 10%, 15.1% and 9.4% respectively. Extremely weak distracters were those that had a negative impact, which means the high-grade students were more attracted than the low-grades. Of the MCQs, 42.7% had one, and one third of the questions had two weak distracters. Among all 642 MCQ questions, only 15.1% had three good distracters.*

Conclusion *Our study showed that more than half of the distracters had a weak quality. So we suggested that before performing a test, the experts' opinion about the designed questions should be obtained and considered to minimize the technical problems of the questions.*

Key Words *MULTIPLE-CHOICE TEST, DISTRACTER, TEST ANALYSIS*

Introduction

Multiple-choice tests (MCTs) are the most useful objective tests that are used to evaluate students' learning level in different educational areas and also to evaluate various capabilities. If designed properly, these tests can evaluate and measure most of the educational outcomes such as reasoning, judgment and prediction skills (1, 2).

Correspondence Maryam-Sadat Kaveh Tabatabaee. Educational Development Center, Mashad University of Medical Sciences, Mashad, Iran.
Tel (+98)-21-511-8412081
Tabatabaee121@hotmail.com

Each question is composed of three parts: the body, the correct answer and a number of distracters. The aim of using distracters is to distract those who do not know the correct answer. In general, a multiple choice question (MCQ) must be designed in a way that only the knowledgeable examinees answer it correctly and others should not be able to find the correct answer. Therefore a distracter must have a reasonable appearance and should seem to be true at first look. In other words, it should not be overtly wrong to attract the unprepared examinees (1, 3-6). However, it is expected that some of the correct answers will be chosen based on estimation and guess and this

possibility increases with decreasing number of the questions in the test, very hard questions, and disregarding the structural principles in designing the questions (1, 6). In general, more attention to the structural principles can lead more know-ledgeable students to find the correct answer, and the role of chance in identifying the correct answer will be minimized (1, 3, 7).

The evaluation of the current tests demon-strates the weak design and structure of the questions (8) and sometimes the questions are equivocal, vague, controversial or trivial (9). Using the question bank of 17 nursing faculty members, Masters et al. examined 2913 MCQs with respect to their structural principles and cognitive domain (10). The results showed that 2233 questions had a somehow weak structure and 47.3% just examined the students' memory, not their reasoning skills. Also, the results of Brozo's study on 1220 MCQs designed by 36 faculty members confirmed the weak structure in teacher-made tests so that 44% of the questions had at least one directive lead or clue and that in 70% of these cases, the students could find the correct answer using these clues. In addition, the results of another study on 300 MCQs in Hamadan Medical School (1999) show that 37% of the questions have at least one or two defects with 32.9% vague questions and lack of rational coordination of distracters in 14.5% of question (11). Another study on the structure of MCQs in Basic Science Comprehensive Exams also confirms these results, so the most common problem in designing an MCQ was the vagueness of the question body and using inappropriate choices (7). Moreover, 71% of examinees in pre-residency exam in Kerman University reported that more than one third of the questions had inappropriate distracters (12).

It should be noted that an inappropriate distracter could act as a directive clue and help the unprepared examinee to find the correct answer. On the other hand, if each choice refers to a separate subject, the question will not be an MCQ anymore and each choice becomes a true-false question, which is independent of other choices (1, 6). In addition, the inappropriate distracters can decrease the validity and reliability of the test (13). Therefore, analysis of the distracters and comparing the frequency of answers in low and high-grade examinees is necessary. Considering the important role of distracters in differentiating between low and high-grade students, this study was performed in Mashad Medical School in 2001-2002 to analyze the quality of distracters in MCTs.

Materials and Methods

In this descriptive study, 932 answer sheets were obtained from the educational department of Mashad Medical School, which were completed by the students in 21 MCTs given in the first semester of 2001-2002. After correcting the sheets in the Medical Educational Development Center and completing the data form of each exam, the data were recorded by Optical Mark Reader (OMR). Then, considering the proportion of questions in each exam and according to the basic, general or specific subject of the exam, 642 MCQs out of 1284 were randomly selected (which included 1926 distracters). After evaluating and comparing the frequency of distracter selection in the groups of low and high-grade students, it was decided whether the distracters had a good efficacy or not, according to the general rule in the analysis of distracters. The general rule in the analysis of distracters is as follows: each distracter should attract at least one person of the low-grade group and if some members of each group are attracted, the number of low-grades must be more than high-grades (1, 6). Therefore, eight types were considered for each distracter that is shown in Table 1.

It should be mentioned that for determining the high and low-grade groups, the scores of the answer sheets were arranged in ascending order, and 27% of the highest scores and 27% of the lowest scores were assigned to the high and low-grade groups, respectively. The data obtained from the answer sheets were analyzed using SPSS software.

Results

According to the data gathered from 21 MCTs in Mashad Medical Schools, 14.28%, 61.9% and 23.8% of the exams were conducted in the course of basic science, pathophysiology and clinical clerkship, respectively. Also, most of the exams subjects (90.48%) were clinical topics. The mean number of questions in each test was 30.6±10.5, and the highest and lowest number of question where given in the kidney disorders exam (10 questions), and Urology and Endocrinology exams (50 questions each), respectively. The average number of examinees studied in all the

TABLE 1 THE TYPE OF DISTRACTER ACCORDING TO THE DISTRIBUTION OF ANSWERS IN THE LOW-GRADE AND HIGH-GRADE GROUPS

	Frequency of choosing the distracter in low- and high-grade groups	
	High-grade	Low-grade
Extremely good	Zero	More than 50%
Very good	Zero	33.3 to 50%
Good	Zero	Less than 33.3%
Relatively Good	The low-grades choose at least 16% more than high-grades	
Relatively weak	The low-grades choose 5-15% more than high-grades	
Weak	Both groups choose equally (the difference is less than 5%)	
Very weak	None of the groups choose the distracter	
Extremely weak	The high-grades choose more than low-grades (negative impact)	

TABLE 2 THE DISTRIBUTION OF MCQS IN MASHAD MEDICAL SCHOOL ACCORDING TO THE NUMBER OF WEAK DISTRACTERS

The number of weak distracter in the question	Questions	
	Number	Percent
Without weak distracter	97	15.1
One weak distracter	274	42.7
Two weak distracters	202	31.5
Three weak distracters	69	10.7
Total	642	100

exams was 44.4±15.2. Pharmacology exam had the lowest number of attendants (26) while Urology exam had the highest number (86).

As figure 1 shows, more than one third (34.5%) of the distracters had a "weak" to "extremely weak" rating. The members of both groups were attracted equally to 10% of the distracters ("weak"). In addition, the high-grade group chose 9.4% of the distracter more than low-grade group, which means that these distracters had a negative impact in student evaluation ("extremely weak"). Among the distracters, 11.4% had a "relatively weak" function.

On the other hand, Table 2 shows that 42.7% of the questions had at least one weak distracter and one third of the questions had two such choices. So in general, only 15.1% of the questions had three good distracters (Figure 2).

Discussion

According to the results of this study, 45.9% of distracters had a "relatively weak" to "extremely weak" function, so these choices must be reconsidered if the questions are to be used again in the future (1). Also, 15.1% of distracters attracted none of the members of low and high-grade groups. These choices are completely useless and their presence or absence has no effect on the results of student evaluation, so these must be replaced with new choices (1). On the other hand, although the aim of using distracters is to separate the prepared and unprepared examinees (1, 6, 9), 9.4% of distracters had a negative impact. In fact, the high-grade students are usually being punished because of their meticulousness in finding the errors and complexities of the questions (1). Vague statements and phrases should not be used in an MCQ (1, 3, 6, 9). Some believe that an MCQ must be vague enough to draw a curtain over the correct answer and convert it into a riddle. However, this will eliminate the very purpose of the evaluation (16).

Considering the fact that increasing the number of good distracters can minimize the role of chance and guessing in selecting the correct answer, and that most studies suggest that three distracters are enough (13), if the structural requirements are not met in designing an MCQ, the question will act as a set of true-false questions. Therefore, the student
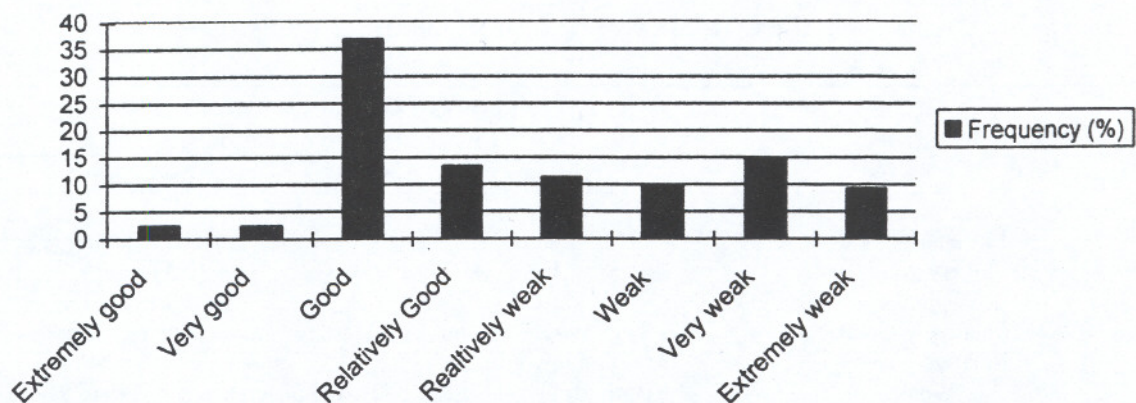
FIGURE 1 THE DISTRIBUTION OF 1926 DISTRACTERS IN MCTs IN
MASHAD MEDICAL SCHOOL ACCORDING TO THEIR FUNCTION.

has to choose between two answers, true or false, and the role of chance will increase to 50%. A clear example is when two choices are contradictory and one of them is the correct answer. In general, the choices must be homogeneous and relevant to the body of the question and there should not be any clue that guides the examinee to the correct answer. It is recommended that the distracters be designed based on the previous exams and according to the common mistakes of the students (3).

It should be emphasized that in order to design better MCQs, the opinions of educational experts and the statistical results of test analyses must be taken into consideration and a question bank should be constructed. A formal educational course for teaching structural principles of an MCT and other evaluation methods to the instructors is also recommended.

Acknowledgments

References

1. Seif AA. Methods of educational measurement and evaluation. Doran Publ. Tehran, Iran, 1999. [Farsi]
2. Killoran J. In defense of the multiple-choice question. Med Educ 1992;56(2):106-8.
3. Dent JA. Harden RM. A practical guide for medical teachers. Churchill-Livingstone 2001.
4. Frey R. Multiple-choice item writing dos and don'ts. 1995 [Online]. Available from: URL:http://www.eric.ed.gov/searchdb/search db.html No. ED398238.
5. Brozo WG. A study of test-wiseness clues in college and university teacher-made tests with implications for academic assistance centers. College reading and learning assistance technical report, 1984:84-101.
6. Hooman HA. Cognitive and educational measurements and the technique of test design. Parsa Publ. Tehran, 10th edition, 1998. [Farsi]
7. Khoshkam M, Shajari J. The structure of questions in the Basic Science Comprehensive Test. J of Research in Med Educ Spring 1998:160.
8. Paxton M. A linguistic perspective on multiple-choice questioning. Assessment & Evaluation in Higher Education 2000;25(2): 109-19.
9. Gilbert GG. An educational booklet for health workers. 1985.
10. Masters JC, Hulsmeyer BS, Pike ME, Leichty KM, Verst AL. Assessment of multiple-choice questions in selected test banks accompanying textbooks used in nursing education. Journal of Nursing Education. 2001;40(1):25-32.
11. Arab M. Assessment of the errors in multiple-choice questions in the tests performed in Hamadan University of Medical Sciences in the second semester of 1999. The 4th symposium on medical education. Journal of Tehran Medical School 2000 Oct:51. [Farsi]
12. Sarrafinejad A, Haghdoost AA. Assessment of the advantages and disadvantages of the content and the method of pre-residency exam

in the viewpoint of examinees in Kerman. The 4th symposium on medical education. Journal of Tehran Medical School 2000 Oct:48 [Farsi]

13. Haladyna TM, Dowing SM. A quantitative review of research on multiple-choice item writing.1985 [Online]. Available from: URL: http://www.eric.ed.gov/searchdb/searchdb/.ht ml No: ED255580.

14. Matlock HS. Basic concepts in item and test analysis. Paper presented at the annual meeting of the Southwest educational research association. 1997 Jan [Online]. Available from: URL:http://www.eric.ed.gov/searchdb/searchdb.html No: ED406441

15. Kehoe J. Writing multiple-choice test items. 1995 [Online]. Available from: URL:http://www.eric.ed.gov/searchdb/searchdb.html No: ED398236

16. Levola S. Measurement and evaluation in education and training. 1998 Summer.