

Item analysis of in use multiple choice questions in pharmacology

Mandeep Kaur, Shweta Singla, Rajiv Mahajan

Department of Pharmacology, Adesh Institute of Medical Sciences and Research, Bathinda, Punjab, India

ABSTRACT

Background: Multiple choice questions (MCQs) are a common method of assessment of medical students. The quality of MCQs is determined by three parameters such as difficulty index (DIF I), discrimination index (DI), and distracter efficiency (DE). **Objectives:** The objective of this study is to assess the quality of MCQs currently in use in pharmacology and discard the MCQs which are not found useful. **Materials and Methods:** A class test of central nervous system unit was conducted in the Department of Pharmacology. This test comprised 50 MCQs/items and 150 distracters. A correct response to an item was awarded one mark with no negative marking for incorrect response. Each item was analyzed for three parameters such as DIF I, DI, and DE. **Results:** DIF of 38 (76%) items was in the acceptable range ($P = 30-70\%$), 11 (22%) items were too easy ($P > 70\%$), and 1 (2%) item was too difficult ($P < 30\%$). DI of 31 (62%) items was excellent ($d > 0.35$), of 12 (24%) items was good ($d = 0.20-0.34$), and of 7 (14%) items was poor ($d < 0.20$). A total of 50 items had 150 distracters. Among these, 27 (18%) were nonfunctional distracters (NFDs) and 123 (82%) were functional distracters. Items with one NFD were 11 and with two NFDs were 8. Based on these parameters, 6 items were discarded, 17 were revised, and 27 were kept for subsequent use. **Conclusion:** Item analysis is a valuable tool as it helps us to retain the valuable MCQs and discard the items which are not useful. It also helps in increasing our skills in test construction and identifies the specific areas of course content which need greater emphasis or clarity.

Key words: Difficulty index, discrimination index, distracter efficacy, item analysis, multiple choice questions

Submission: 29-12-2015 **Accepted:** 18-04-2016

INTRODUCTION

Multiple choice questions (MCQs)/items are the most common method of assessing the knowledge capabilities of undergraduate, graduate, and postgraduate students in medical colleges. These can be used for both formative and summative assessments. Framing of good MCQs is a time-consuming and a challenging process. It is said that

appropriately constructed MCQs result in objective testing that can measure knowledge, comprehension, application, analysis, and evaluation.^[1] Hence, MCQs to be used must be of quality and they need to be tested for the standard or quality. Item analysis is one such tool that provides information regarding the reliability and validity of a test items. Item analysis examines the student responses to individual test items/MCQs to assess the quality of those items and test as a whole.^[2] In simple terms, it is a process of collecting, summarizing, and using information from students' responses to assess the quality of test items. It includes three parameters such

Address for correspondence: Dr. Mandeep Kaur,
Adesh Institute of Medical Sciences and Research, Flat No. 307,
B Block, Bathinda, Punjab, India.
E-mail: mandeeplobana1@yahoo.com

Access this article online	
Quick Response Code:	Website: www.ijabmr.org
	DOI: 10.4103/2229-516X.186965

This is an open access article distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 License, which allows others to remix, tweak, and build upon the work non-commercially, as long as the author is credited and the new creations are licensed under the identical terms.

For reprints contact: reprints@medknow.com

How to cite this article: Kaur M, Singla S, Mahajan R. Item analysis of in use multiple choice questions in pharmacology. Int J App Basic Med Res 2016;6:170-3.

as difficulty index (DIF I), discrimination index (DI), and distracter efficiency (DE).^[3]

Keeping in view the increasing importance of MCQs in various medical examinations and entrance tests, this study was undertaken to assess the quality of MCQs/items currently in use in pharmacology and discard the MCQs which are not found useful.

MATERIALS AND METHODS

A class test was held in the Pharmacology department on central nervous system in which 150 students appeared. A total of 50 MCQs or items and 150 distracters were analyzed. Each MCQ comprised a stem and four responses, and the students selected one best answer from these four choices. A correct response to an item was awarded one mark and the wrong one zero, there was no negative marking. To avoid possible cheating of answers from neighboring student, four sets of question papers were made with disorganized sequencing of questions.

After evaluation of class test, marks obtained by the students were arranged in descending order and entered in Microsoft office excel sheet 2007. The upper one-third students (50) were considered as high achievers (H) and lower one-third (50) as low achievers (L). For computation purpose, marks obtained by middle one-third were discarded. Each item was analyzed for three parameters.^[3]

Difficulty index/facility value/*P* value

It is the percentage of students in high and low achievers group who answered the item correctly. It ranges between 0% and 100%. It was calculated using the formula $DIF I \text{ or } P = \frac{H + L}{N} \times 100\%$; where, *H* = number of students answering the item correctly in the high achieving group, *L* = number of students answering the item correctly in the low achieving group, and *N* = total number of students in the two groups (including nonresponders). Results of DIF I was interpreted as shown in Table 1.

Discrimination index or *d* value

DI is the ability of an item to differentiate between students of higher and lower abilities and ranges between 0 and 1. It was calculated using the formula $DI = \frac{2 \times (H - L)}{N}$ where, the symbols *H*, *L*, and *N* represent the same values as mentioned above. Results of DI were interpreted as shown in Table 2.

Distracter efficiency

DE is determined for each item on the basis of the number of nonfunctional distracter (NFD) (option selected by <5% of students) in it. Results of DE were interpreted as shown in Table 3.

RESULTS

After statistical analysis, it was found out that DIF I of 38 (76%) items was in the acceptable range, 11 (22%) items were too easy, and 1 (2%) item was too difficult [Figure 1].

DI of 31 (62%) items was excellent, of 12 (24%) items was good, and of 7 (14%) items was poor [Figure 2].

The present study showed that there were 123 functional distracters and 27 NFDs out of total 150 distracters in 50 MCQs [Table 4].

DISCUSSION

The tests incorporating MCQs are commonly used method of assessing the cognitive domain of learning, though psychomotor and affective domains cannot be assessed.^[4] Still, it has an advantage of testing large number of students in a short time period with quick and easy marking. An appropriately

Table 1: Interpretation of difficulty index

Cutoff points (%)	Interpretation	Action
<30	Too difficult	Revise/discard
30-70	Good/acceptable	Store
50-60	Excellent/ideal	Store
>70	Too easy	Revise/discard

Table 2: Interpretation of discrimination index

Cutoff points	Interpretation	Action
<0.20	Poor	Discard
0.2-0.34	Good	Revise
≥0.35	Excellent	Store

Table 3: Interpretation of distracter efficiency

NFDs	DE (%)	Action
3	0	Discard
2	33.3	Revise
1	66.6	Store
0	100	Store

NFDs: Nonfunctional distracters; DE: Distracter efficiency

Table 4: Distracter efficiency of multiple choice questions

Indicators	Values
Number of items	50
Distracters	150
Functional distracters (%)	123 (82)
NFD (%)	27 (18)
Items with 1 NFD (DE=66.6%)	11
Items with 2 NFDs (DE=33.3%)	8
Item with 3 NFDs (DE=0%)	0
Item with no NFDs (DE=100%)	31

NFD: Nonfunctional distracter; DE: Distracter efficiency

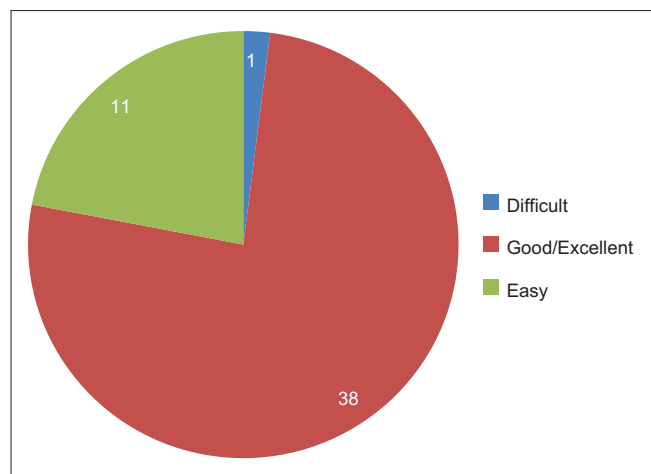


Figure 1: Difficulty index of multiple choice questions

constructed and framed MCQ needs to be tested for the standard or quality. Item analysis is one such tool which is a valuable yet relatively simple procedure performed after the examination that provides information regarding the reliability and validity of a test.^[5] It is of great help in improving the quality of items and prepares a viable question bank for subsequent use. It is also helpful to both students and teachers as it provides feedback to the teacher to improve their method of teaching and encourage the learners to learn more effectively.^[6]

The parameter DIF is a misnomer as more is the DIF I, easier is the question and vice versa, so it is also termed as easy index or facility value (FV) by few authors.^[7,8] It helps in determining whether the students learned the concept being tested.

In a study conducted by Patil and Patil^[9] on 100 MBBS students of medicine for 100 MCQs, mean DIF I of 48.90 ± 13.72 was reported. In this study, the *P* value of 35 (22%) items was in the acceptable range (30–70%), 25 (25%) items was ideal (50–60%), 18 (18%) items was too easy ($P > 70\%$), and 22 (35%) items was too difficult ($P < 30\%$). In another study on item analysis done by Patel and Mahajan^[10] on 150 MBBS students for MCQs test with 50 questions, 10 (20%) items were in unacceptable range ($P < 30\%$ or $P > 70\%$) and 40 (80%) items were in acceptable range ($P = 30\text{--}70\%$). Item analysis done by Mehta and Mokhasi^[3] on 100 MBBS students for MCQs test comprising 50 questions in the subject of anatomy reported mean DIF I of 63.06 ± 18.95 with DIF I of 31 (62%) items in the acceptable range ($P = 30\text{--}70\%$), 16 (32%) items were too easy ($P > 70\%$), and 3 (6%) items were too difficult ($P < 30\%$). Kolte^[11] reported mean DIF I as 57.92 ± 19.58 . In this study, the *P* value of 26 (65%) items was in acceptable range (30–70%), 10 (25%) items were easy ($P > 70\%$), and 4 (10%) items were difficult ($P < 30\%$). Our findings corresponded with the study

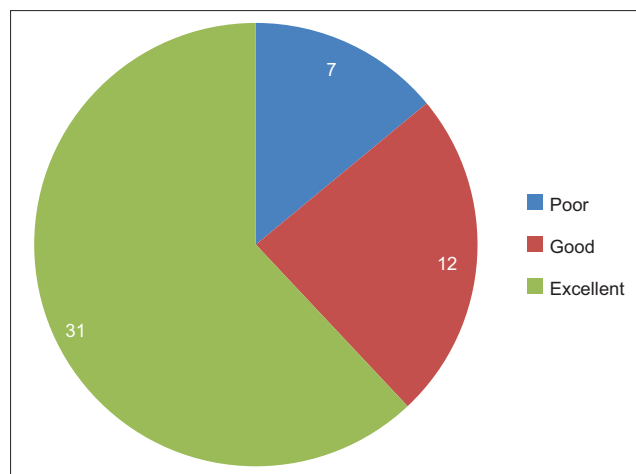


Figure 2: Discrimination index of multiple choice questions

by Mehta and Kolte having a mean DIF I of 59.18 ± 15.14 . The *P* value of 38 (76%) items was in the acceptable range ($P = 30\text{--}70\%$), 11 (22%) items were too easy ($P > 70\%$), and 1 (2%) item was too difficult ($P < 30\%$). Too difficult items (DIF I $\leq 30\%$) can lead to deflated scores, while the easy items (DIF I $> 70\%$) may result into the inflated scores and a decline in motivation.^[8] Items with high DIF I ($> 70\%$) should be placed either at the start of the test as “warm-up” questions to boost the confidence of students or discarded, similarly items with low DIF I ($< 30\%$) should be either revised or removed altogether. Our study had only one item which was too difficult and it was discarded. Items which were too easy were 11 and these were revised and kept for subsequent use along with items within acceptable range.

DI is another important parameter of item analysis that helps in detecting the ability of items to discriminate between skilled and unskilled examinees. DI normally ranges from 0 and 1, but sometimes its value can be < 0 when this index is called negative DI.^[8] It is because of more number of students from lower achiever group able to answer the item correctly in comparison with students from high achiever group. The reason for this negative value is either due to ambiguous question or an answer key that was wrongly marked. Our study items did not show negative DI. In an item analysis study by Patil and Patil,^[9] out of total 100 items, 24 had DI < 0.2 (poor), 45 had DI ≥ 0.20 and ≤ 0.35 (good), and 31 had DI > 0.35 (excellent). A study by Singh et al.^[12] on item analysis of 20 MCQs reported 6 items (30%) with DI < 0.2 , 4 (20%) items with DI ≥ 0.20 and ≤ 0.35 , and 10 (50%) items with DI > 0.35 . In another study by Patel and Mahajan^[10] on item analysis of 50 items, 9 items had DI < 0.2 , 21 items had DI ≥ 0.20 , and ≤ 0.35 and 20 items had DI > 0.35 . In a study by Mehta and Mokhasi^[3] on item analysis, mean DI was 0.33 ± 0.18 . Out of total 50 items, 15 (30%) items had DI < 0.2 , 9 (18%) items was DI ≥ 0.20 and ≤ 0.35 , and 26 (52%) items

had $DI > 0.35$. Our study was in accordance with this study and showed mean DI of 0.37 ± 0.15 with 7 (14%) items had $DI < 0.2$, 12 (24%) items had $DI \geq 0.20$ and ≤ 0.35 , and 31 (62%) items had $DI > 0.35$. Seven items with $DI < 0.2$ were discarded altogether due to their poor discriminating power. Twelve items with $DI \geq 0.20$ and ≤ 0.35 were revised and kept for subsequent use along with items with high discriminating power ($DI > 0.35$).

The cardinal rule during framing of good MCQs is that the distracters must be plausible, i.e., closely placed to the correct answer. This will increase the chances of choosing these distracters over the correct answer by the learners. Implausible distracters deny chances to test a learner. In a study conducted on 50 items with 150 distracters by Gajjar *et al.*,^[8] 133 were functional distracters and 17 were NFDs with an overall mean DE of 88.6 ± 18.6 . Mehta and Mokhasi^[3] had reported that in total 150 distracters, 53 were NFD, 28 were functional distracters, and 69 had none response with mean DE of 63.97 ± 33.56 . In a study of item analysis by Patil and Patil^[9] on 100 items, 263 were functional distracters and 37 were NFDs with an overall mean DE of 82.8 ± 15.6 . Our study showed mean DE of 83.98 ± 24.52 , with 123 functional distracters and 27 NFDs. None of the items had three NFDs. Items with no NFDs were 31 ($DE = 100\%$).

It is concluded from the present study that considerable test items were within the recommended values by experts. However, some test items did not meet the requirement of well-designed question items. Hence, these items can be revised or discarded, and a viable question bank can be prepared.

Financial support and sponsorship

Nil.

Conflicts of interest

There are no conflicts of interest.

REFERENCES

1. Kemp JE, Morrison GR, Ross SM. Developing evaluation instruments. In: Designing Effective Instruction. New York: MacMillan College Publishing Company; 1994. p. 180-213.
2. Sharif MR, Rahimi SM, Rajabi M, Sayyah M. Computer software application in item analysis of exams in a college of medicine. *J Sci Eng Technol* 2014;4:565-9.
3. Mehta G, Mokhasi V. Item analysis of multiple choice questions – An assessment of the assessment tool. *Int J Health Sci Res* 2014;4:197-202.
4. Singh T, Anshu. Principles of Assessment in Medical Education. New Delhi: Jaypee Brothers Medical Publishers; 2012. p. 89.
5. Considine J, Botti M, Thomas S. Design, format, validity and reliability of multiple choice questions for use in nursing research and education. *Collegian* 2005;12:19-24.
6. Pande SS, Pande SR, Parate VR, Nikam AP, Agrekar SH. Correlation between difficulty & discrimination indices of MCQs in formative exam in physiology. *South East Asian J Med Educ* 2013;7:45-50.
7. Singh T, Gupta P, Singh D. Principles of Medical Education. 4th ed. New Delhi: Jaypee Brothers Medical Publishers; 2013. p. 109.
8. 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 Community Med* 2014;39:17-20.
9. Patil VC, Patil HV. Item analysis of medicine multiple choice questions (MCQs) for under graduate (3rd year MBBS) students. *Res J Pharm Biol Chem Sci* 2015;6:1242-51.
10. Patel KA, Mahajan NR. Itemized analysis of questions of multiple choice question exam. *Int J Sci Res* 2013;2:279-80.
11. Kolte V. Item analysis of multiple choice questions in physiology examination. *Indian J Basic Appl Med Res* 2015;4:320-6.
12. Singh JP, Kariwal P, Gupta SB, Shrotriya VP. Improving multiple choice questions (MCQs) through item analysis: An assessment of the assessment tool. *Int J Sci Appl Res* 2012;1:53-7.