

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

Vijaya K. Suryadevara*, Zaheda Bano

Department of Pharmacology,
Guntur Medical College,
Guntur, Andhra Pradesh, India

Received: 12 July 2018

Accepted: 16 July 2018

***Correspondence to:**

Dr. Vijaya K. Suryadevara,
Email: drsvkdrsvk@gmail.com

Copyright: © the author(s), publisher and licensee Medip Academy. This is an open-access article distributed under the terms of the Creative Commons Attribution Non-Commercial License, which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

ABSTRACT

Background: In medical education, multiple choice questions/Items are the more frequently used assessment tools to assess the knowledge abilities and skills of medical students, for being their objectivity, wide coverage in less time. However only the Quality Items gives a valid and reliable assessment. The quality of an Item is determined by difficulty index (DIF I), Discrimination Index (DI) and Distractor efficiency (DE). Aim of the study was to know the quality of Items in pharmacology by Item analysis and to develop a MCQs bank with quality Items. **Methods:** The present study was conducted on 150 II MBBS students of Guntur Medical College, AP, India. A class test containing 50 Items with 150 distractors from topic chemotherapy was conducted. Item with the correct choice/response was awarded with one mark and with the wrong choice zero marks, no negative marks. Each test Item was analysed with DIF I, DI and DE and the results were tabulated and tested statistically, with unpaired "t" test.

Results: Mean DIF I, DI, DE values with standard deviations in the present study are 44.72+17.63%, 0.30+0.12%, 84.48+24.65 respectively. DIF I of 32 (64%) items was good to excellent range (31%-60%) 9 (18%) Items were easy (>61%) and 9(18%) Items were difficult (>30%). DI of 10 (20%) Items was good (0.15 to 0.24.) 29 (58%) Items were excellent with DI > 0.25 and 11 (22%) Items were poor with DI <0.15. Among 150 distractors, 127 (85%) were functional distractors (FDs) and 23 (15%) were non-functional distractors (NFDs). DE of 33 (66%) items with nil NFDs was 100%, for 12 (24%) Items with one NFD, was 66.6%, for 4 (8%) items with 2 NFDs was 33.3% and for 1 (2%) Item with 3NFDs DE was 0%. When unpaired "t" test was applied to the means of "difficult" and "easy" Items, 96.22+11.33% SD, 51.44+29.31% SD respectively, the p-value obtained was 0.00058, which was highly significant.

Conclusions: The study showed that Item analysis is a valid tool to identify quality Items, which assess, the students' knowledge abilities and discriminate different levels of performance abilities of students effectively.

Keywords: Assessment, Difficulty index, Discrimination index, Distractor efficiency, Item analysis, Non-functional distractor

INTRODUCTION

Medical students acquire medical knowledge and skills by various learning practices. During and at the end of the learning process, assessment of the student's knowledge and skills is very crucial. Hence a valid tool is necessary,

to know the insight about their learning and competencies. Among the various assessment tools, "Multiple Choice Questions" (MCQs), also called as "Items" are frequently used for assessment of the students, because of their objectivity, wide coverage of subject in less time and minimal assessor's bias.¹ Appropriately constructed MCQs

are useful to measure not only knowledge but also comprehension, application, analysis, and evaluation.²

Usually, an Item (MCQ) consists of a stem with four choices/responses. Among the four choices one choice is correct and the remaining three are incorrect, called as "distractors" as they distract the student from choosing the correct choice. An Items test is said to be valid to assess the student, when it contains quality Items only. Whereas the quality of an Item is determined by three indices, Difficulty Index (DIF I) also denoted by FV (facility value) or P-value, Discriminatory Index (DI) and Distractor Efficiency (DE).^{3,4}

As Items are widely used for assessing the students, the present study was undertaken with the aim to identify the quality Items in Pharmacology with the help of Difficulty Index (DIF I), Discriminatory Index (DI) and Distractor Efficiency (DE) and to develop a Pharmacology MCQs/Items bank with quality Items.

METHODS

In the present study, II MBBS 5th semester students of Guntur Medical College, Guntur, Andhra-Pradesh, India, were included. After 30 hours of didactic lectures on chemotherapy and giving ample time for preparation, a class test in Pharmacology on chapter Chemotherapy was conducted for these students. The test was conducted under strict vigilance with the help of other faculty in the department and also optimum distance was maintained between the students to avoid possible copying from the neighbouring students.

Out of 173 students 150 students attended the test. The test contained 50 Items and each Item was with a single stem and four choices/ responses. Among four choices, one choice was correct, and the remaining three choices (distractors) were incorrect. The students were asked to select and round off only one correct choice to each Item. Later, 50 Items along with 150 distractors in each of 150 test papers were evaluated. An Item with correct response was awarded with one mark, whereas zero marks were awarded to an Item with wrong response, without any negative marks. After awarding the marks to all the 150 students, marks were entered in descending order in Microsoft office Excel version 2007, so that the highest marks were placed as first one and the least marks as the last one in the marks list. Then, the list was divided into 3 equal groups with 50 each. The upper one-third (50) higher marks, which were considered as high ability (H) group and the lower one-third (50) lesser marks which were considered as low ability (L) group were taken for computation purposes. The middle one-third (50) medium marks were not taken, assuming that they behave in similar pattern. For each Item in all (100) test papers of both high ability group and low ability group students, Difficulty Index, Discrimination Index and Distractor Efficiency were calculated with the help of following formulae:

$$\text{DIF I} = \frac{[H + L]}{N} \times 100$$

$$\text{DI} = 2 \times \frac{[H - L]}{N}$$

Here N = total number of students in both high and low ability groups and H and L are number of correct responses in high and low ability groups, respectively.^{2,5,6}

DIF I of an Item ranges between 0% and 100%. It indicates the percentage of students who answered the Item correctly. Higher the value of DIF I, Item is an easy one and it can be answered by most of the students. If DIF I value is low, Item is said to be a difficult one. DIF I of an ideal Item ranges between 41% and 60%.

DI of an Item ranges between 0 and 1. It describes the ability of an Item to differentiate between students of higher and lower abilities. If DI value is high for an Item, it discriminates more effectively, the students of higher and lower abilities. DI of an ideal Item is 1 and it perfectly discriminates the students of lower and higher abilities.

DE of an Item ranges between 0% and 100%. It indicates the number of non-functional distractors (NFDs) in an Item. If any choice (other than correct choice) of an Item is selected by <5% of students, the choice is called as "NFD" and any choice (other than correct choice) of an Item is selected by 5% or more than 5% of students, it is called as functional distractor "FD". If an Item has 3 NFDs, the DE of that Item is 0%, or 2NFDs, DE will be 33.3%, or one NFD, DE will be 66.6% and if it has nil NFDs then DE will be 100%.^{2,7,8}

After calculating DIF I, DI and DE for all 50 Items in each 100 test papers of high and low ability groups, the results were tabulated. To obtain the significance/p value of the study results statistically, unpaired "t" test was applied.

RESULTS

Out of 50 marks the maximum and minimum marks scored by the students were 44 and 8 respectively.

Table 1: Mean with standard deviation of various indices of Items in the test (N= 50).

Parameter	Mean±Standard deviation
Difficulty Index % (DIF I)	44.72±17.63
Discrimination Index (D I)	0.30±0.12
Distractor Index % (D E)	84.48±24.65

For 50 Items in the test, Mean Difficulty Index was 44.72%±17.63 SD, Mean Discrimination Index was 0.30±0.12 SD and Mean Distractor Efficiency was 84.48%±24.65 SD (Table 1).

DIF I of 20 (40%) Items was "excellent" with a range of 41%-60% and DIF I of 12 (24%) Items was "good" with a

range of 31%-40%. Whereas 9 (18%) Items were "easy" with a DIF I of >61% and another 9 (18%) Items were "difficult" with a DIF I of <30% (Table 2).

Table 2: Distribution of Items in relation to DIF I and actions proposed.

DIF I cut off %	Items (N=50)	Interpretation	Action
≤30	9	Difficult	Revise/discard
31-40	12	Good	Store
41-60	20	Excellent	Store
≥61	9	Easy	Revise/discard

DI of 10 (20%) Items was "good" with a range of 0.15 to 0.24. Whereas 29 (58%) Items were "excellent" with DI ≥0.25 and 11(22%) Items were "poor" with DI <0.15 (Table 3).

Table 3: Distribution of Items in relation to D I and actions proposed.

D I Cut off points	Items (N=50)	Interpretation	Action
<0.15	11	Poor	Discard/revise
0.15 - 0.24	10	Good	Store
≥0.25	29	Excellent	Store

The test contained 50 items with 150 distractors. Among 150 distractors, 127 (85%) were functional distractors (FDs) and 23 (15%) were non-functional distractors (NFDs). DE was 100% for 33 (66%) Items with nil NFDs and it was 66.6% for 12 (24%) Items with one NFD each.

Whereas DE was 33.3% for 4 (8%) Items with 2 NFDs each and it was 0% for 1 (2%) Item with 3NFDs (Table 4).

Among 17 Items with NFDs, DIF I of 8 (47%) Items was "easy" and for 7 (41%) Items it was "excellent". Whereas DIF I for 1 (6%) Item was "good" and for 1 (6%) Item it was "difficult". DI of 10 (59%) Items was "excellent", for 4 (23%) Items it was "good" and for 3 (18%) Items it was "poor" (Table 5).

Table 4: Distractor analysis and Distractor Efficiency.

Number of Items	50
Total Distractors	150
Functional Distractors (FDs)	127
Non Functional Distractors (NFDs)	23
No. of Items with "0" NFDs (DE=100%)	33
No. of Items with "1" NFDs (DE=66.6%)	12
No. of Items with "2" NFDs (DE=33.3%)	4
No. of Items with "3" NFDs (DE=0 %)	1

Table 5: Items with Non Functional Distractors and their relationship with DIF I and D I.

DIF I (%)	Items with NFDs (N = 17)	D I	Items with NFDs (N = 17)
≤30 (Difficult)	1	<0.15 (Poor)	3
31-40 (Good)	1	0.15-0.24 (Good)	4
41-60 (Excellent)	7	≥0.25 (Excellent)	10
≥61 (Easy)	8		

Table 6: Distractor Efficiency of Items with different values of DIF I and D I.

Grading	Difficulty Index (DIF I)		Discrimination Index (DI)	
	Difficult (≤30%)	Easy (≥61%)	Poor (<0.15)	Excellent (≥0.25)
No of items	9	9	11	29
DE (%) Mean±SD	96.22±11.33	51.44±29.31	90.73±15.88	84.86±23.08
Unpaired t - test	0.00058		0.444157	

Table 7: Comparing data of similar studies of Item analysis with the present study.

Indices	Gajjar and Sharma ²	Kaur and Singla ⁶	Arora and Prithi ¹¹	Patil and Patil ¹²	Mehta and Mokhasi ¹³	Present study
DIF I Mean±SD%	39.4±21.4	59.18±15.14	44.8±17.13	48.90±13.72	63.06±18.95	44.72±17.63
Easy (%)	10	22	Nil	18	32	18
Acceptable (%)	24	76	75	60	62	64
Difficult (%)	16	2	25	22	6	18
DI Mean±SD	0.14±0.19	0.37±0.15	0.37±0.18	0.19±0.13	0.33±0.18	0.30±0.12
Poor	52	14	15	24	30	22
Good	18	24	25	45	18	20
Excellent	30	62	60	31	52	58
DE Mean±SD%	88.60±18.60	83.98±24.52	87.5±17.2	82.8±15.6.	63.97±33.56	84.48±24.65

The Mean DE of 9 Items with DIF I as "difficult" was $96.22 \pm 11.33\%$ SD and the Mean DE of 9 Items with DIF I as "easy" was $51.44 \pm 29.31\%$ SD. When unpaired "t" test was applied to these Means of "difficult" and "easy" Items, the value obtained was 0.00058. The Mean DI of 11 Items with DI as "poor" was 90.73 ± 15.88 SD and the Mean DI of 29 Items with DI as "excellent" was 84.86 ± 23.08 SD. When unpaired "t" test was applied to these Means of "poor" and "excellent" Items, the value obtained was 0.44 (Table 6).

DISCUSSION

Item analysis is a valuable, relatively simple procedure performed after the test, that provides information regarding the quality of test Items.⁹ Nowadays, the most common type of Items employed in examinations are type A MCQs, which consist of a stem followed by four or five options or distractors.¹⁰ In the present study, a class test containing 50 type A Items was conducted for 150 II MBBS students. When these Items were analysed, the mean DIF I with SD was $44.72 \pm 17.63\%$, which was within excellent range (41%-60%) and corresponded with the mean DIF I with SD of $44.8 \pm 17.13\%$, reported in a study conducted by Ardra and Prithi on 120 II MBBS students with 20 MCQs in microbiology (Table 7).¹¹

Out of 50 test Items in this study DIF I of 32 (64%) Items was in acceptable range of 31%-60% and were stored as MCQs Question Bank for future use, whereas 9 (18%) Items were easy ($\geq 61\%$) and 9 (18%) Items were difficult (≤ 30). This study findings were nearer to another study conducted by Patil and Patil with 100 MCQs in medicine for MBBS students (Table 7).¹² Too difficult Items ($\geq 30\%$) may give deflated scores and too easy Items (DIF I $> 70\%$) may give inflated scores and less motivation.¹ The 9 easy Items in the study were slightly revised and kept for future use as warm up questions, to boost the confidence of students and 9 difficult Items were thoroughly checked for possible confusing language, areas of controversy, for any incorrect key and after revision, they were kept to develop MCQs bank.⁵

In Item analysis, DI helps to know the ability of Items to discriminate between students of higher and lower abilities and normally ranges from 0 to 1. In the present study, mean DI was 0.30 ± 0.12 SD with excellent discriminating power and nearer to the value of mean DI 0.33 ± 0.18 SD reported in a study conducted by Mehta and Mokhasi with a test of 50 Items in anatomy on 100 first MBBS students (Table 7).¹³ Out of 50 Items, in this study, 11(22%) Items were with poor DI (< 0.15), 10 (20%) Items were with good DI (0.15-0.24) and 29 (58%) Items were with excellent DI (≥ 0.25). A total of 39 (78%) Items were with good to excellent (0.15-1) discriminating power and were stored as MCQs bank for further use and 11 (22%) Items with poor DI values, were revised and stored. These DI values of present study, when compared with an Item analysis study by Patil and Patil reported that similar percent (76%) of

Items had good to excellent discriminating power.¹² Another study by Mehta and Mokhasi¹³ also showed 70% percent of Items with well discriminating power nearer to our study value (78%) (Table 7).

Apart from the DIF I, DI, the quality of an Item is determined by another important index called DE. While framing good quality Items, the cardinal rule is that the distractors must be plausible, i.e., nearer to the correct answer/choice.⁶ Otherwise an implausible distractor deny chances to test a student. In the present study, the 50 test Items had a total of 150 Distractors, among which 127 distractors were FDs, and remaining 23 were NFDs. Our study mean DE with SD value $84.48 \pm 24.65\%$ was nearer to the mean DE value of $83.98 \pm 24.52\%$ reported in a study conducted by Kaur and Singla⁶ with 50 Test Items in pharmacology on II MBBS students. When compared to another study of Item analysis by Gajjar and Sharma² with 50 test Items and 150 distractors in community medicine on 148 first MBBS students, a similar mean DE value of $88.60 \pm 18.60\%$ was observed (Table 7). The DE value of an Item depends on the presence or absence of NFDs in that Item. When an Item is with more NFDs, they increase DIF I i.e., makes an Item easy and reduces DE value of the Item, conversely an Item is with more FDs, they decrease DIF I i.e., makes an Item difficult and increases DE value of the Item. In this study, 33 (66%) Items had zero NFDs with a DE of 100%, which were retained for MCQs bank, 12 (24%) Items had 1NFD with a DE of 66.6%, which were revised and retained, 4 (8%) items had 2NFDs with a DE of 33.3% which were discarded and 1 (2%) Item had 3NFDs with a DE of 0%, so it was discarded. The difference between mean DE of difficult Items ($96.22 \pm 11.33\%$) and the mean DE of easy Items ($51.44 \pm 29.31\%$) was statistically highly significant as the p-value was 0.00058, when unpaired t test was applied (Table - 6). It shows a strong relation between DE and DIF I. While the difference between the mean DE of Items with poor DI (90.73 ± 15.88) and the mean DE of Items with excellent DI (84.86 ± 23.08) was statistically not significant as the p-value was 0.444157, when unpaired t test was applied.

CONCLUSION

Nowadays test Items are used more frequently as an assessment tool while assessing student's learning abilities and skills. For a valid and reliable assessment, quality Items are required. The conducted study showed nearly 2/3 of the 50 test Items were quality Items and some Items needed revision, and a very few Items were discarded. It concludes that Item analysis is a very useful and necessary procedure to obtain valid Items from time to time and to develop a MCQs Bank for future use.

ACKNOWLEDGEMENTS

Authors would like to thank their Institution, II MBBS students, Pharmacology faculty of Guntur Medical

College, Guntur, AP, India, for allowing them and for supporting to conduct this study.

Funding: No funding sources

Conflict of interest: None declared

Ethical approval: The study was approved by the Institutional Ethics Committee (GMC/IEC/112/2017)

REFERENCES

1. 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 Journal of Community Medicine: Official Publication of Indian Association of Preventive and Social Medicine.* 2014;39(1):17-20.
2. Kemp JE, Morrison GR, Ross SM. Developing evaluation instruments. *Designing Effective Instruction.* New York: MacMillan College Publishing Company; 1994:180-213.
3. Sarin YK, Khurana M, Natu MV, Thomas AG, Singh T. Item analysis of published MCQs. *Indian Pediatr.* 1998;35:1103-5. (PubMed:10216545).
4. 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;9:1-8. (PMCID:PMC2713226) (PubMed: 19580681)
5. Hingorjo MR, Jaleel F. Analysis of one-best MCQs: The difficulty index, discrimination index and distracter efficiency. *J Pak Med Assoc.* 2012;62(2):142-7.
6. Kaur M, Singla S, Mahajan R. Item analysis of in use multiple choice questions in pharmacology. *Int J App Bas Med Resear.* 2016;6(3):170-3.
7. Singh T, Gupta P, Singh D. Test and item analysis. *Principles of Medical Education.* 3rd ed. New Delhi, Jaypee Brothers Medical Publishers (P) Ltd; 2009:70-77.
8. Namdeo SK, Sahoo B. Item analysis of multiple choice questions from an assessment of medical students in Bhubaneswar, India. *Int J Res Med Scienc.* 2016 Dec 30;4(5):1716-9.
9. Sidine 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.
10. Skakun EN, Nanson EM, Kling S, Taylor WC. A preliminary investigation of three types of multiple choice questions. *Med Educ.* 1979;13:91-6 [PubMed].
11. Menon AR, Kannambra PN. Item Analysis to Identify Quality Multiple Choice Questions. *NJLM.* 2017;6(2):MO07-10.
12. Patil VC, Patil HV. Item analysis of medicine multiple choice questions (MCQs) for under graduate (3rd yr MBBS) students. *Res J Pharm Biol Chem Sci.* 2015;6:1242-51.
13. 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.

Cite this article as: 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. *Int J Basic Clin Pharmacol* 2018;7:1517-21.