Original article:

The Correlation between the Acceptable Range of Difficulty and Discrimination Indices in Four-Response Type Multiple Choice Questions in Physiology

[1]Dr. Manish Badkur , [2]Dr.Garima Suryavanshi* , [3]Dr.Ashwin K Abraham

[1]M.D. Anaesthesia, Intensivist, National cancer institute, Nagpur

[2]M.D. Physiology, Assistant Professor, Dept of Physiology, Indira Gandhi Medical College,Nagpur

[3]Resident, Dept of Physiology, Indira Gandhi Medical College, Nagpur

Corresponding author *

Abstract:

Introduction: The competency assessment of doctors both in terms of knowledge & clinical skills by multiple choice questions plays a vital role in various examinations. The standard indices which are used to evaluate the competency of multiple choice questions (MCQs) in various examinations are Difficulty index (P) and discrimination index (D).

Objectives: The study was undertaken to find out correlation between Difficulty index (P) and discrimination index (D).

Materials and Methods: 150 1st year M.B.B.S students were subjected to 60 MCQ taken from the entire syllabus of Physiology as part of their periodic assessment. The Pearson correlation between Difficulty index (P) and discrimination index (D) was calculated using graph pad prism version 7.03.

Results:Both the indices showed wide variation. Item difficulty index ranged from 13% to 93% while discrimination index ranged from 0.5 to 0. In terms of difficulty index out of 60 MCQs about 44 MCQs (73%) were in acceptable range (P= 30-70%) while 8 MCQs (13%) were too easy (P >70%), 8 MCQs (13%) were too difficult (P<30%). In terms of discrimination index 28 MCQs (47%) were in good range (D = 0.2 to 0.35),15 MCQs (25%) were in excellent range (D >0.35) & 17 MCQs (28%) were in poor range (D < 0.2). So in total 43MCQs were in acceptable range. The correlation between acceptable range discrimination & difficulty index was significant. The correlation of MCQs in acceptable range discrimination index with too easy & too difficult MCQs was non-significant.

Conclusion:The majority (73%) of the items were acceptable as far as difficulty and discriminative indices were concerned. The correlation between discrimination index of acceptable range & moderately easy/difficult items was significant, while too easy and too difficult items had non significant correlation with acceptable range discrimination index.

Keywords - Item Analysis, Difficulty index , Discrimination index.

Introduction

 In medical field the competency of doctors involves assessment of 3 important 'domains'- cognitive, psychomotor and affective. So medical assessment programmes must be designed to assess the above mentioned domains. (1) But the current medical assessment programmes still cannot test all the 3 domains together at the same time. (2) Objective evaluation in medical field by the means of match the following, true or false statements and fill in the blanks can be used for assessing the cognitive domain. However, MCQs are the most popular form of objective evaluation. (3) One-best MCQ is an efficient tool to evaluate higher-order cognitive domain of Bloom's taxonomy such as interpretation, synthesis and application of knowledge. It can also assess progress of the student by identifying their strengths and weaknesses and can provide feedback to teachers regarding class performance. (4,5,6)

Formulating ideal MCQs is a complex, challenging, tedious and time consuming process in a multidisciplinary integrated curriculum. One of the major aspect in the formation of an MCQ or a test item is its reliability & validity which in turn is determined by Classical Test Theory (CT) item analysis. (7,8)

CT Item analysis is the simple procedure of collection, summarisation and using information from students' responses to assess the quality of test items after the end of examination. (7,8). The 2important item characteristics in Classical Test Theory (CT) item analysis arei) difficulty index, and ii) discrimination index which tests the reliability of different testitems. (9) The difficulty index is the percentage of learners who answered an item correctly and ranges from 0 to 100%. Closer the value of difficulty index of an item to 0%, the more is the item difficult while value approaching more towards 100%, the item is easier. Thus it identifies whether the item is too difficult or too easy & finally helps in differentiating between students who have learned & understood the content and those who have not. Items with difficulty index between 30-70% are considered acceptable. (10)

While the discrimination index of an item discriminates between students who are high or low scorers. It ranges from -1 to +1. The closer this value is to +1, the excellent is the discrimination index of an item indicating high scoring students select the correct answer for each item more often than the low- scoring students. While if, the low scoring students got a specific item correct more often than the high scoring ones, then that item has a value approaching towards -1. (6,10) So analysing both difficulty and discrimination indices of each item provides information regarding what the students have learned and enables teachers to determine and correct the faulty items. (10) Thus to find out how these 2 indices correlatethe following study was undertaken.

Objectives

To find out correlation between Difficulty index (P) and discrimination index (D) of individual test item.

Material And Methods

The study was done in the Department of Physiology, Indira Gandhi Govt. Medical College, Nagpur as a part of regular Periodic assessment of students. Hundred & fifty First-year MBBS students were voluntarily involved in the study. They were subjected to 60 MCQ questions with single best response covering entire syllabus of Physiology. The MCQs' were constructed to assess various levels of knowledge according to Bloom's Taxonomy (11). Formulation of MCQs' was done by the Head of Department and other Professors in the department. There was no negative marking and the time allotted was sixty minutes. Evaluation was done out of sixty marks and the Difficulty & discrimination index of all 60 MCQs were calculated. Then the correlation was found between the two. (12).

Steps in item Analysis (13,14)

The scores of all the students were arranged in descending order of merit. Then they were divided into three groups as low, moderate & high achievers in accordance to their marks. Top one third students were considered as high achievers H [n=50] and bottom one third group as low achiever's L [n=50]. Each item then was analysed for the Difficulty & discrimination index & correlation was found between them.

1. Calculation of Difficulty Index

    P value was calculated using the formula

    $$P = H + L / N \times 100$$

Where

- H= number of students answering the item correctly in high achievers group.
- L= number of students answering the item correctly in the low achievers group.
- N= Total number of students in the two groups (including non-responders).

P value (%)	Item Interpretation
< 30 %	Too Difficult
30 – 70 %	Acceptable
50 - 60 %	**Ideal**
>70 %	Too Easy

Items having P value less than 30% & more than 70% are not acceptable and need modification.

2. Calculation of **Discrimination index**

$$D = H-L \times 2/N$$

Where the symbols H, L and N are same as above.

D value	Item Interpretation
< 0.20	Poor
0.20-0.35	Good
>0.35	**Excellent**
>0.20	**Acceptable**

Items with D value less than 0.20 are not acceptable & needs revision.

3. All the MCQs with discrimination index in acceptable range were not same as that of MCQs with difficulty index in acceptable range. So the correlation was found using graph pad prism version 7.03

   i. Between MCQs which are common in acceptable range of discrimination & difficulty index.

   ii. Between MCQs with discrimination index in acceptable range with difficulty index either too easy or too difficult

   iii. Between MCQs with poor discrimination index with difficulty index in acceptable range.

Results

Both the indices showed wide variation. Item difficulty index ranged from 13% to 93% while discrimination index from 0.5 to 0. In terms of difficulty index out of 60 MCQs about 44 MCQs (73%) were in acceptable range (P= 30-70%) while 8 MCQs (13%) were too easy (P >70%), 8 MCQs (13%) were too difficult (P<30%) hence 16 items that means 26.66% could be used after modification. [Table1].

While in terms of discrimination index 28 MCQs (47%) were in good range (D = 0.2 to 0.35), 15 MCQs (25%) were in excellent range (D >0.35) & 17 MCQs (28%) were in poor range (D < 0.2). So a total of 43 MCQ (28+15) were in acceptable range in terms of discrimination index. [Table2].

There were 43 & 44 MCQs in acceptable range in terms of discrimination & difficulty index respectively. [Table1 & 2]. Out of 43 MCQs, 30 MCQs were common in acceptable range of both the indices. While there were 7 MCQs with discrimination index in acceptable range with difficulty index either too easy or too difficult. There were 6 MCQs with poor discrimination index with difficulty index in acceptable range. 1 MCQ of acceptable difficulty index had poor discrimination index. 30 common MCQs of acceptable range showed significant correlation (r = 0.978, p=<0.0001) between discrimination & difficulty index. While the correlation was non-significantbetween 7 MCQs (r = 0.288, p = 0.5312) with discrimination index in acceptable range & difficulty index either too easy or too difficult and 7 MCQs (r = -0.2715, p =0.5558) with difficulty index in acceptable range & poor discrimination index.[Table3]

Table 1: Difficulty Index (P) of items analysed.

DI Range	Number of items	Interpretation	Action taken
30 to 70 %	**44(73.33%)**	**Acceptable**	**In MCQ bank**
>70 %	8(13.33%)	Too easy	Revise
<30 %	8(13.33%)	Too difficult	Revise

Table 2: Discrimination index(D) of items analyzed

Range(D)	Number of items	Interpretation	Action taken
0.2 to 0.35	28	Good	In MCQ bank
>0.35	15	Excellent	In MCQ bank
>0.2	**43**	**Acceptable**	**In MCQ bank**
<0.2	17	Poor	Revise

Table 3: Correlation between Discrimination index (D) & Difficulty Index (P) of items analyzed

Indices in Acceptable Range.	No. Of MCQs	rvalue	pvalue	interpretation	Action taken
Both D & P	30	0.978	<0.0001	Significant	In MCQ bank
Only D	7	0.288	0.5312	Non-Significant	Revise
Only P	7	-0.2715	0.5558	Non-Significant	Revise

Discussion

Developing the perfect, flawless assessment test seems to be an unattainable goal for mostteachers. In light of this fact,the calculation of difficulty and discrimination index provides a valuable tool in designing the test. These indices enable teachers to review and improve the whole process of evaluation by guiding them how effective the test questions are in assessing the knowledge of students. (15) So the present study was undertaken to find correlation between discrimination and difficulty index.

In our study out of 60 items only 44 items (73.33%) in terms of difficult index were in acceptable range (P = 30-70%) while 8 MCQs (13%) were too easy (P >70%), 8 MCQs (13%) were too difficult (P<30%). This is probably due to poor understanding of difficult topics, ambiguity in wordings of the questions

or even inappropriate key and may also be due to personal variations in students' intelligence. Hence 16 items that means 26.66% could be used after modification. While 43 items (71.66%) in terms of discrimination index were in acceptable range (D > 0.2) & the remaining 17 MCQs (28%) were in poor range (D <0.2). As the discrimination index (D) serves as an effective tool regarding quality of each item. So Items with poor discrimination should be reviewed &needs modification (15).

Similar results were depicted in other studies like Mitra, N.K et al. (7) & Sim et al (5).We also calculated correlation between the 2 indices. We found that 30 common MCQs of acceptable range showed significant correlation (r = 0.978, p=<0.0001) between discrimination & difficulty index. While the correlation was non-significantbetween 7 MCQs (r = 0.288, p = 0.5312) with discrimination index in acceptable range & difficulty index either too easy or too difficult and 7 MCQs (r = -0.2715, p =0.5558) with difficulty index in acceptable range & poor discrimination index.Thus we can infer that relationship between difficulty and discrimination indices was not linear. Maximal discrimination occurred with moderatelyeasy/difficult items. Too easyand too difficult items showed poordiscrimination.Similar observation was reported by Si-Mui Simet al., (5) in their study.So too difficult and too easy items needed further modification.

Conclusion

In our study the majority of items 44(73%) fulfilled the criteria of acceptable difficulty and Discrimination index. But the correlation between 2 indices was significant for only 30MCQs (68%). So total out of 60 MCQs, only 30 (50%) MCQs can be added to departmental question bank. Thus we recommend not only estimation of indices but also finding correlation between them to develop the most authentic question bank possible for assessment of knowledge & skills.

References

1. Karelia BN, Pillai A, Vegada BN. The levels of difficulty and discrimination indices and relationship between them in four-response type multiple choice questions of pharmacology summative tests of Year II M.B.B.S students. IeJSME 2013; 7(2): 41-46.

2. Ho TF, Yip WCL, Tay JSH. The use of multiple choice questions in medical examination: An evaluation of scoring and analysis of results. Singapore Med J 1981; 22(6): 361-7.

3. Ananthakrishnan N. Item analysis-validation and banking of MCQs. In: Ananthakrishnan N, Sethuraman KR, Kumar S, editors. Medical Education principles and practice. 2nd ed. JIPMER, Pondicherry: 131-7.

4. De Champlain AF, Melnick D, Scoles P, Subhiyah R, HoltzmanK,Swanson D, et al. Assessing medical students' clinical sciences knowledge in France: a collaboration between the NBME and a consortium of French medical schools. Acad Med 2003;78: 509-17.

5. Sim S M, Rasiah RI. Relationship between Item Difficulty and Discrimination Indices in True/False-Type Multiple Choice Questions of a Para-clinical Multidisciplinary Paper. Ann Acad Med Singapore 2006;35:67-71.

6. Hingorjo RM, Jaleel F.Analysis of One-Best MCQs: the Difficulty Index, Discrimination Index and Distractor Efficiency. J Pak Med Assoc. February 2012 ;6(2):142-147.

7. Mitra N K, Nagaraja H S, Ponnudurai G, Judson J P.The Levels Of Difficulty And Discrimination Indices In Type A Multiple Choice Questions Of Pre-clinical Semester 1 Multidisciplinary Summative Tests. IeJSME 2009;3(1):2-7.

8. Zubairi AM, Kassim NLA. Classical and Rasch analysis of dichotomously scored reading comprehension test items. Malaysian J of ELT Res. 2006; 2: 1-20.

9. Siri A, Freddano M. The use of item analysis for the improvement of objective examinations. Procedia - Social and Behavioral Sciences. ICEEPSY 2011(29):188 – 197.

10. KOÇDAR S, KARADAĞ N, Murat Doğan ŞAHIN MD.Analysis of the Difficulty and Discrimination Indices of Multiple-Choice Questions According to Cognitive Levels in an Open and Distance Learning Context. TOJET: The Turkish Online Journal of Educational Technology – October 2016;15(4):16-24

11. Bloom, B., Englehart, M. Furst, E., Hill, W., &Krathwohl, D. (1956). Taxonomy of educational objectives: The classification of educational goals. Handbook I: Cognitive domain. New York, Toronto: Longmans, Green.

12. Kashyap S.Item Analysis Of Multiple Choice Questions. International Journal of Current Research.December, 2015;7(12):24057-24058.

13. EVRORO, Sylvanus E. ITEM ANALYSIS OF TEST OF NUMBER OPERATIONS.Asian Journal of Educational Research.2015; Vol. 3(1):18-25.

14. Boopathiraj C, Chellamani K. Analysis Of Test Items On Difficulty Level And Discrimination Index In The Test For Research In Education.*International Journal of Social Science & Interdisciplinary Research*. FEBRUARY 2013;Vol.2 (2):189-193.

15. Pande SS, Pande SR, Parate VR, Nikam AP, Agrekar SH. Correlation between difficulty and discrimination indices of MCQ's in formative exam in physiology. South East Asian J Med Educ2013;7:45-50.