Original Article

Item analysis: A tool to increase MCQ validity

[1]Dr. Minu Ramakrishnan, [2]Dr. Aditya B. Sathe, [3]Dr. Vinayak A. Sabnis

[1]Professor and Head, Department of Ophthalmology, K J Somaiya Medical College, Mumbai.

[2]Professor, Department of Biochemistry, K J Somaiya Medical College, Mumbai.

[3]Professor, Department of Pathology, K J Somaiya Medical College, Mumbai.

Corresponding author: Dr. Minu Ramakrishnan

Abstract

Background: Multiple Choice Questions (MCQs) are a popular assessment tool due to their high objectivity, high reliability and the ability to assess a large content in a short time span. However their validity cannot be taken for granted and needs to be ascertained by various means. Item analysis is one such process of testing MCQ validity by assigning certain numeric indices to them and comparing these with acceptable standards.

Aims and objectives: The aim of the present study was to demonstrate the use of item analysis as a tool to ascertain MCQ validity.

Material and methods: Sixty MCQs in ophthalmology of the single best response type were subjected to the process of item analysis and the difficulty index, discrimination index and distracter effectiveness were calculated using standard formulae. These MCQs and distracters were then classified into groups as per standard reference ranges of these parameters and the absolute number and percentage of MCQs and distracters in each group were calculated.

Results: The percentage of MCQs that could be accepted as having desired validity based on difficulty index, discrimination index and distracter effectiveness were 50%, 33% and 32% respectively while 63% of the distracters were functional (acceptable). Thus these MCQs could be used for assessment while the rest needed modification and retesting or needed to be discarded.

Conclusion: MCQ item analysis must be performed to filter MCQs of acceptable validity which would increase their quality as assessment tools thereby making assessment more meaningful.

Key words : MCQ, item analysis, difficulty index, discrimination index, distracter effectiveness.

Introduction

Assessment is an integral part of curriculum and is used to guide future learning (formative assessment) or to judge competence to practice (summative assessment).[1] Multiple choice questions (MCQs) are a widely used tool in assessment protocols. MCQs have the advantage of having a high degree of objectivity and reliability and can assess a large area of the content in a small time span. [1,2] Although objectivity and reliability are inherent qualities of MCQs their validity cannot be assumed due to the possibility of the 'student guessing the right answer without knowing it'. Medical education technology recommends the implementation of standard prevalidation and postvalidation protocols to increase the validity of MCQs. While prevalidation prevents errors in framing MCQs by using guidelines and checklists, post validation helps to identify MCQs with questionable validity so that they can be appropriately modified before reuse or discarded.

www.ijbamr.com P ISSN: 2250-284X , E ISSN : 2250-2858

Item analysis is a postvalidation procedure that characterizes every MCQ and its distracters by assigning a numerical value to it in the form of a difficulty index, a discrimination index and a distracter effectiveness. Based on standard acceptable limits of these indices MCQs can be either accepted for banking, or modified and revalidated or discarded.

The objective of this study was to perform an item analysis of MCQs for testing their validity as assessment tools.

Materials and methods

The study was conducted at a medical college in the city of Mumbai. This is a retrospective observational study using records of three MCQ based formative examinations conducted in the year 2015-16. A total of 60 MCQs in the subject of ophthalmology used in these three examinations (20 MCQs in each examination) were included in the study. All MCQs were of the single best response type with one right answer and three distracters. A total of 50 students had attempted these MCQs.

Students were listed in the decreasing order of marks obtained in the respective examination. The upper 17 and lower 17 students formed the high achiever group and the low achiever group respectively. Based on the answers marked by the students the difficulty index, discrimination index and the distracter effectiveness were calculated for each MCQ and distracter using standard formulae mentioned in table 1. These MCQs and distracters were then classified using cut off values into groups as mentioned in table 1.[2]

Table 1: Formulae for item analysis and categories of classification of MCQs and distracters based on difficulty index, discrimination index and distracter effectiveness.

Parameter	Difficulty index	Discrimination Index	Distracter effectiveness
Formula	$\frac{(H+L)}{T} \times 100$	$\frac{(H-L)}{T} \times 2$	Percentage of students having marked that distracter as the right answer
Categories and cutoffs	Very Difficult: Difficulty index less than 30%.	Good discriminator: Discrimination index more than or equal to 0.2	Functional Distracter (FD): Distracter effectiveness more than or equal to 5%
	Acceptable: Difficulty index 30% to 70% ::	Poor Discriminator: Discrimination index less than 0.2	Non functional Distracter (NFD): Distracter effectiveness less than 5%:
	Very Easy: Difficulty index above 70%:		

H: Number of High achievers who have answered the question correctly

www.ijbamr.com P ISSN: 2250-284X , E ISSN : 2250-2858

L: Number of Low achievers who have answered the question correctly

T: Total number of students considered for analysis

Results

Results of data analysis for difficulty index, discrimination index and distracter effectiveness are presented in Figure 1, Figure 2 and Figure 3 and figure 4.

Figure 1: Classification of MCQs (n=60) as per difficulty index and the number and percentage of MCQs in each class

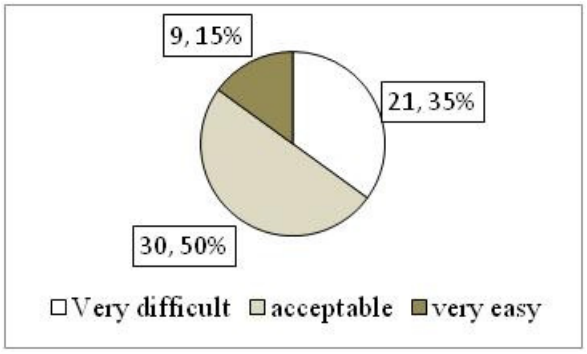


Figure 2: Classification of MCQs (n=60) as per discrimination index and the number and percentage of MCQs in each class

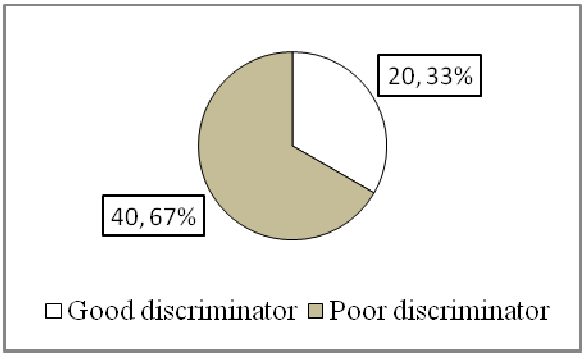


Figure 3: Classification of MCQs (n=60) as per distracter effectiveness and number and percentage of MCQs in each class

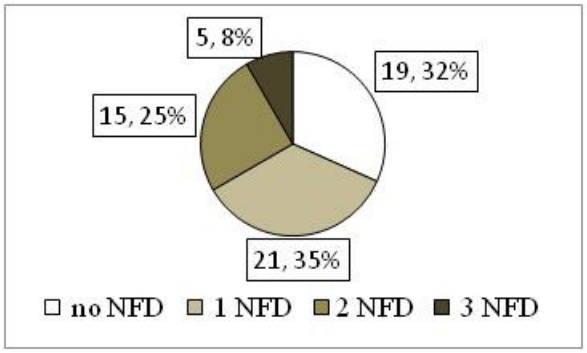
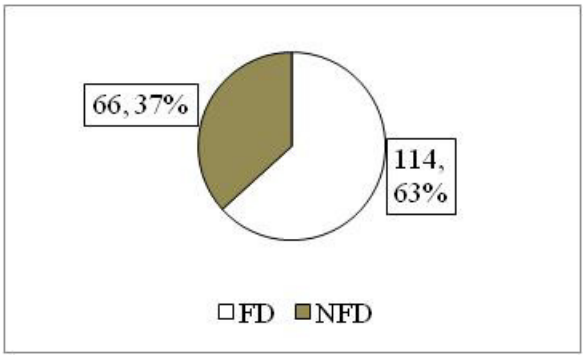


Figure 4: Classification of distracters (n=180) as per distracter effectiveness and number and percentage of distracters in each class



Discussion

The present study was conducted with the objective of implementing the procedure of item analysis of MCQs to determine the validity of MCQs and distracters. Results revealed that the percentage of MCQs that were in the acceptable range based on difficulty index, discrimination index and distracter effectiveness were 50%, 33% and 32% respectively while 63% of the distracters were functional (acceptable). Thus these MCQs and distracters can be added to the question bank while the rest have to modified or replaced and retested until they satisfy the criteria of acceptability.

In a study by Halikar S et al item analysis of twenty MCQs in ophthalmology was performed. Results showed that the percentage of acceptable MCQs based on difficulty index and discrimination index were 35%, 50% respectively. All MCQs had at least one non-functional distracter. The percentage of functional distracters in this study was found to be 23%. The authors concluded that item analysis can generate a bank of validated MCQs with known values of indices from which question paper setters can choose the appropriate MCQs for a given examination.[3] Namdeo SK et al performed an item analysis of 25 MCQs in pediatrics and reported that 60% and 68% of MCQs were acceptable based on

difficulty index and discrimination index respectively. 12%MCQs had no Nonfunctional distracters. 46% of the distracters were found to be functional. The authors concluded that item analysis is helpful to determine technical faults in MCQs and provides valuable information to modify them appropriately to increase their validity.[4] An item analysis performed on 50 MCQs in anatomy by Mehta G et al revealed that 62% and 70% of MCQs were in the acceptable range of difficulty index and discrimination index respectively. 34% of MCQs had no NFDs and 18% distracters were functional. The authors concluded that item analysis is vital for developing MCQs with a high pedagogic and psychometric value.[5]

In addition to the standard interpretation of MCQs mentioned in table 1, item analysis can identify flaws in MCQs. While as per standard interpretation a difficulty index more than 70% or less than 30% indicates a too easy and too difficult question respectively, a flawed MCQ could also yield these values. Some flaws in MCQs could provide a clue to the answer making the question easier while flaws resulting in ambiguity in the question may make the MCQ more difficult.[6] Such an MCQ will also have a low discrimination index since it will affect the performance of high achievers adversely while inflating the performance of low achievers.[6] These flaws could be in the stem, in the key or in the distracters and can be identified and rectified by content experts. Most of these flaws can be rectified during prevalidation but those that escape can be picked up during item analysis and rectified. Thus item analysis is also a valuable means to identify flaws in MCQs.

Despite MCQ item analysis being a valuable tool, many medical college departments may not voluntarily adopt it due to lack of awareness, lack of compulsion from regulatory bodies, time and labor involved and a perception by teachers that their subjective validation may be sufficient and equivalent to the objective item analysis procedure. However a subjective validation is extremely variable from teacher to teacher and its sensitivity in validation of MCQs is relatively low as compared to the standard item analysis procedure.[7] Moreover, use of software tools can significantly reduce time and labor involved in item analysis. Thus creating awareness, providing software support and establishment of a clear mandate by regulators can popularize the procedure of item analysis increasing the validity of assessment.

The less number of students in high and low achiever groups (seventeen each) can be cited as a limitation of this study. However it does not undermine the importance of the conclusion drawn regarding importance of item analysis that has also been proved by similar studies cited above.

To conclude, item analysis is a useful tool to establish the validity of MCQs. A large proportion of MCQs framed by teachers may not be acceptable or valid. Such MCQs can be identified by item analysis and suitably modified before being used as assessment tools. Item analysis can also identify flawed MCQs that can be suitably modified to increase its validity thus helping to achieve the overall assessment objective.

www.ijbamr.com P ISSN: 2250-284X , E ISSN : 2250-2858

References

1. Epstein RM. Assessment in Medical Education. N Engl J Med 2007;356:387-96

2. Rege NN, Validation of MCQs. In: Bhuiyan PS, Rege NN, Supe AN, eds The art of teaching medical students 2002; 2nd Ed Medical Education Technology cell, Seth GSMC and KEM hospital:239-51.

3. Halikar S, Godbole V, Chaudhari S. Item Analysis to Assess Quality of MCQs. Medical Science 2016;6:123-5

4. Namdeo SK, Sahoo S. Item analysis of multiple choice questions from an assessment of medical students in Bhubaneswar, India. Int J Res Med Sci 2016;4:1716-9.

5. 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.

6. Omer AA, Abdulrahim ME, Albalawi IA. Flawed. Multiple-choice questions put on the scale: What is their impact on students' achievement in a final undergraduate surgical examination? J Health Spec 2016;4:270-5.

7. Sathe AB, Bhadre RB, Bhalkar MS, Mosamkar SP. Subjective versus objective item analysis of Multiple Choice Questions. International J. of Healthcare and Biomedical Research 2017;5:89-94.