

## Research Article

# Using item analysis on essay types questions given in summative examination of medical college students: facility value, discrimination index

Gul-Ar Navi Khan<sup>1\*</sup>, Nazia Ishrat<sup>2</sup>, Abdul Qayyum Khan<sup>3</sup>

<sup>1</sup>Department of Physiology, J.N.M.C., AMU, Aligarh, U.P., India

<sup>2</sup>Department of Obstetrics & Gynaecology, J.N.M.C., AMU, Aligarh, U.P., India

<sup>3</sup>Department of Orthopaedics, J.N.M.C., AMU, Aligarh, U.P., India

**Received:** 19 November 2014

**Accepted:** 8 December 2014

### \*Correspondence:

Dr. Gul-Ar Navi Khan,

E-mail: dr.gular@rediffmail.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:** The most common tool used for assessment of knowledge is the essay questions. Their evaluations depend upon test and item analysis which is consisting of analysis of individual questions and analysis of the whole test. The objective of our study was to calculate the Item analysis (Facility value, FV and discrimination index, DI) of questions given in terminal examination of MBBS students and observes the adequacy of questions framing in examination.

**Methods:** Study contains 150 medical students undergone terminal examination consist of questions of essay type, structured essay type and short answer questions were given to the students. We divide them into high ability group and low ability group. Mark range was decided for essay type, structured essay type and short answer type questions and FV & DI was calculated.

**Results:** Studies described that facility values of 62.5 percentage questions were come under recommended & acceptable range and 50 percentage questions were come under acceptable range. Discrimination value of 100 percentage questions were come under recommended & acceptable range where 75 percentage questions were come under acceptable range.

**Conclusion:** The importance of item analysis is highlighted from our results. For improvement of examination items with average difficulty and high discrimination should be implemented into future examinations to improve test scores and properly discriminate among the students.

**Keywords:** Item analysis of essay type questions, Facility value, Discrimination index

## INTRODUCTION

The majority of faculty member of any academic institution evaluate their students every year. Evaluation is the integral part of teaching learning process and it is of many types depending upon the curriculum of the institution. In the evaluation process, framing of questions should be such that this evaluation process become reliable and valid. The academic institutions evaluate their student by various types of questions like

essay type, modified essay type, short answer type and multiple choice questions. Depending upon the type of assessment, these questions are framed. Among these questions Multiple Choice Questions (MCQ) are very common and preferred type used in many educational setting. MCQ's are difficult to frame but easy to evaluate because of some known software are available to calculate it. Besides these MCQs questions some other types of questions are also framed and generally given in summative examinations. Few studies demonstrated that

some faculties of college level are unable to frame questions<sup>1-3</sup> and their effects of this poor examination system reflecting in students' performance.<sup>4,5</sup> Since in summative examination (mid-term exam or terminal examination) marks are awarded and added in final totaling so it has high stake value. Such an important and valuable tool for the assessment of teaching and learning of medical students should not be neglected so questions in this context should be properly framed and evaluated whether they are really work for which they are intended to form.

The most common tool used for assessment of knowledge is the essay questions. Their evaluations depend upon test and item analysis which is consisting of analysis of individual questions and analysis of the whole test. Although this activity is more precisely could be done in objective type questions but we have tried it for essay type, structured essay type, and short answer questions. The teaching and learning process is followed by examination. Item analysis is meant for the understanding level of students by evaluating the questions given in examination. It also provides feedback to the examiner also.<sup>6</sup>

In every medical college, examiner believes that their framed questions are satisfactory and they are able to distinguish the real ability of students which reflect the understanding of the material or skill taught and identify whether they are ready proceed in the next stage of teaching or learning.<sup>7</sup>

Item analysis is a valuable procedure performed after the examination that provides information regarding the reliability and validity of a test item.<sup>8</sup>

It also tells how difficult or easy the questions were, calculated by the facility value (difficulty index) and whether the questions were able to discriminate between students who performed well on the test, from those who did not, the discrimination index.<sup>9</sup>

Facility Value (FV) - It is number in the group answering a question right. FV can be calculated by the formula:

$$FV = \frac{HAG + LAG}{N} \times 100$$

HAG - Higher ability group

LAG - Lower ability group

N - Total no. of considered students

Facility value is a measure of how easy or how difficult a question which is given to students so it is also called as Difficulty Index. Higher the FV, easier is the question.

FV value is expressed in percentage. Its range is 0-100.

Its recommended value is 45-60 and its acceptable value is 25-75.

Discrimination Index (DI) - This index indicates the ability of a question to discriminate between a higher and a lower ability student. This is calculated by the formula:

$$DI = \frac{2 \times (HAG - LAG)}{N}$$

DI value is expressed in as a fraction. Its range is 0-1.0.

Its maximum value is 1.0, which indicates an ideal question with perfect discrimination between HAG and LAG. Its value could extend from -1.00 to + 1.00. This minus value is called as negative discrimination which means that more students in the lower group are answering that item correctly than students in the higher group.

Recommended value is >0.25

Acceptable with revision is 0.15-0.25

Discard the question <0.15

These item analyses helps in detecting specific technical flaws in the questions and provide information for improvement. It also increases the skill of examiners in item writing. It enhances further discussion in class and ultimately improving the learning teaching process. A good item is one question, which about half the class can answer i.e., FV of 50%. FV is a useful tool testing adequacy of classroom teaching. A good values of FV suggest that subject area is well known to student, has been well taught, questions has been properly framed. Besides the FV, a good item is one of which Discrimination Index (DI) value is 0.35 or more than 0.35.<sup>10</sup>

The objective of our study was to calculate the Item analysis (facility value and discrimination index) of questions given in terminal examination of MBBS students of 2013 batch and observes the adequacy of questions framing in examination.

## METHODS

Our study contains 150 medical students of batch 2013. A terminal examination was conducted after taking verbal permission from chairman in the department of physiology, J. N. medical college, AMU, Aligarh. Eight questions of essay type, structured essay type and short answer questions were given to the students. Two hours was the time limit and questions paper have 05 marks for individual questions. The copies were evaluated by the individual teacher of respective questions and marks were given. All the medical student's marks were listed for individual question and according to aggregate marks scored, students were arranged in rank order with highest

score given on the top. We divide them into High Ability Group (HAG) and Low Ability Group (LAG). High ability group consist of top ranker 25 students and low ability group consist of low ranker 25 students. Mark range was decided for essay type, structured essay type and short answer type questions with the help of ANGEL group who formulated the score range and mean of individual question.<sup>11</sup> For each question students achieve 5 to 3.5 marks will be considered as correct answer i.e., A. For each question students achieve 3 to 2 marks will be considered as near to correct answer i.e., B. For each question students achieve 1.5 to 0.5 marks will be considered as near to incorrect answer i.e., C. For each question students achieve 0 marks will be considered as incorrect answer i.e., D. For each question, count the no. of students obtained marks of A, B, C, D category.

After we have available data and information about all the given questions in terminal examination of medical student of 2013 batch, we will calculate the indices Facility Value (FV) and Discrimination Index (DI).

### RESULTS

Our study on 150 medical students of batch 2013 of J. N. medical college, AMU, Aligarh. Eight questions of essay type, structured essay type and short answer questions were given to the students in their terminal examination conducted in department of Physiology. We studied Item analysis of individual question and mark range was decided for essay type, structured essay type and short answer type questions with help of ANGEL group who formulated score range and mean of individual question.

**Table 1: For all given questions, no. of student obtained marks in different range is as follows.**

Marks range	5.0-3.5	3.0-2.0	1.5-0.5	0	Total no. of considered students	
Designated sign	A	B	C	D		
No. of HAG students	Q-1	11	14	0	0	25
	Q-2	15	9	1	0	25
	Q-3	5	14	5	1	25
	Q-4	16	9	0	0	25
	Q-5	8	10	7	0	25
	Q-6	22	3	0	0	25
	Q-7	16	8	1	0	25
	Q-8	4	20	1	0	25
No. of LAG students	Q-1	1	3	12	9	25
	Q-2	1	22	2	0	25
	Q-3	0	5	15	5	25
	Q-4	2	13	8	2	25
	Q-5	0	2	20	3	25
	Q-6	4	18	3	0	25
	Q-7	3	14	3	5	25
	Q-8	0	12	11	2	25
<b>Level of correctness</b>	Correct answer	Near to correct answer	Near to incorrect answer	Incorrect answer		

**Table 2: Showing facility value and discrimination index of different questions.**

Question No.	Facility value (FV)		Discrimination index (DI)	
	Value	Range	Value	Range
1	24	Acceptable	0.40	Recommended
2	32	Acceptable	0.56	Recommended
3	10	Not acceptable	0.20	Acceptable
4	36	Acceptable	0.56	Recommended
5	16	Not acceptable	0.32	Recommended
6	52	Recommended	0.72	Recommended
7	38	Acceptable	0.52	Recommended
8	08	Non acceptable	0.16	Acceptable

**Table 3: Showing the Range of Item analysis of questions given in terminal exam.**

Item analysis	Recommended		Acceptable		Non acceptable	
	No. of question	% of question	No. of question	% of question	No. of question	% of question
Facility value (FV)	1	12.5	4	50	3	37.5
Discrimination index (DI)	6	75	2	25	0	0

According to marks range of individual questions, students were counted of both the group and facility value & discrimination index were calculated. Studies described that facility values of 62.5 percentage questions were come under recommended & acceptable range and 50 percentage questions were come under acceptable range while 37.5 percentage questions were come under non-acceptable range.

Discrimination value of 100 percentage questions were come under recommended & acceptable range where 75 percentage questions were come under acceptable range and none of questions was in non-acceptable range.

## DISCUSSION

Development of competent medical personnel's depends upon the knowledge, skills they acquired during their medical teaching. Any assessment whether formative or summative generally taken in medical college education setting, has intense effects on learning. Single correct response type MCQ is an efficient tool for assessment but we have studied the essay type, structured essay type and short answer type. This types of questions are frequently given in examination but lesser studied. In MCQs we have single correct response but in essay type or theoretical questions response is not absolute, given in theoretical ways so its assessment is not so easy. There are very scanty studies of marks distribution on essay type, structured essay type and short answer type questions but marks ranges in our study was formulated with the help of ANGEL 7.4 Instructor Reference Manual made By ANGEL support centre of Eastern Florida state college<sup>11</sup> where the score range and mean of individual question was given and they viewed the response data for Essay and Short answer questions, along with data on how students performed against the question as a whole.

In our study the facility value of maximum no. of questions fall under recommended & acceptable range showed that questions given in terminal examination were of moderately easy type. This condition indicates that most of the given questions need not to be revised. Their value implies that individual teacher did cover the subject matter moderately or the students did show enough interest to the study.

The other objective of our study, the discrimination index of all the questions were came into recommended & acceptable range depicted that all the questions have good

capacity to discriminate between the two group i.e., high ability and low ability group. In our study no item was seen with negative discrimination which reflects that our questions given in terminal examination making true contribution to the evaluation. These items should not be revised since they discriminate between the test takers.

Since given questions were of essay type, structured essay type and short answer type questions so probability of weak students to give answer on guessing basis was eliminated.

The importance of item analysis is highlighted from our results. For improvement of examination items with average difficulty and high discrimination should be implemented into future examinations to improve test scores and properly discriminate among the students.<sup>12</sup>

It might be possible that some of the questions of the tests have facility value and discrimination index less than the recommended or acceptable range so these tests need careful re-examination.

Despite the lack of clear cut guidelines in formulating item analysis gives difficulty to the instructor but regular exercise over this analysis would contribute to us in formulation of appropriate questions. A proper feedback is very helpful to the instructor. These exercises related to items analysis should be regularly analysed and implemented in different level of examinations in medical colleges.

We hope our study would initiate the change the way and support in the designing further method of assessment used in our undergraduate medical curriculum. Fowell and co-workers highlighted the importance of evaluating assessment system where they suggested the role of medical educators in curriculum planning and development where assessment & learning and assessment drive learning become a key part of effective sustainable curriculum development.<sup>13</sup>

## CONCLUSION

The most common tool used for assessment of knowledge is the essay questions. Their evaluations depend upon test and item analysis which is consisting of analysis of individual questions and analysis of the whole test. Although this activity is more precisely could be done in objective type questions but we have tried it for essay

type, structured essay type, and short answer questions. The teaching and learning process is followed by examination. Item analysis is meant for the understanding level of students by evaluating the questions given in examination. These item analyses helps in detecting specific technical flaws in the questions and provide information for improvement. It also increases the skill of examiners in item writing. It enhances further discussion in class and ultimately improving the learning teaching process. The importance of item analysis is highlighted from our results. For improvement of examination items with average difficulty and high discrimination should be implemented into future examinations to improve test scores and properly discriminate among the students. Despite the lack of clear cut guidelines in formulating item analysis gives difficulty to the instructor but regular exercise over this analysis would contribute to us in formulation of appropriate questions. We hope our study would initiate the change the way and support in the designing further method of assessment used in our undergraduate medical curriculum.

#### ACKNOWLEDGEMENTS

We are thankful to all the students of MBBS 2013 batch for their support. We are also thankful to department of physiology & obstetrics gynaecology for helping us in our study.

*Funding: No funding sources*

*Conflict of interest: None declared*

*Ethical approval: Not required*

#### REFERENCES

1. Guthrie DS. Faculty goals and methods of instruction: approaches to classroom assessment. In: Guthrie DS, eds. Assessment and Curriculum Reform. New Directions for Higher Education No. 80. San Francisco: Jossey-Bass; 1992: 69-80.
2. Lederhouse JE, Lower JM. Testing college professor's tests. Coll Student J. 1974;8(1):68-70.
3. McDougall D. College faculty's use of objective tests: State-of-the-practice versus state-of-the-art. J Res Devel Educ. 1997;30(3):183-93.
4. Crooks TJ. The impact of classroom evaluation practices on students. Rev Edu Res. 1988;58(4):438-81.
5. Shifflett B, Phibbs K, Sage M. Attitudes toward collegiate level classroom testing. Educ Res Qrtly. 1997;21(1):15-26.
6. Ebel RL, Frisbie DA. Essentials of educational measurement. In: Ebel RL, Frisbie DA, eds. A Book. 5th ed. Englewood Cliffs, NJ: Prentice-Hall; 1991.
7. Hopkins KD. Educational and psychological measurement and evaluation. In: Hopkins KD, eds. A Book. 7th ed. Needham Heights, MA: Allyn and Bacon; 1998.
8. 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.
9. Sim SM, 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.
10. Singh Tejinder, Gupta Piyush, Singh Daljit. Test and item analysis. In: Singh Tejinder, Gupta Piyush, Singh Daljit, eds. Principles of Medical Education. 3rd ed. Delhi: JayPee Publications; 2009: 70-77.
11. ANGEL group, ANGEL 7.4. Instructor reference manual made by ANGEL support centre of Eastern Florida state college, 2014. Available at: [http://help.angel.msu.edu/sites/default/files/faculty/online\\_documentation\\_74/question\\_pool\\_submission\\_detail.htm](http://help.angel.msu.edu/sites/default/files/faculty/online_documentation_74/question_pool_submission_detail.htm).
12. Wallach PM, Crespo LM, Holtzman KZ, Galbraith RM, Swanson DB. Use of a committee review process to improve the quality of course examinations. Adv Health Sci Educ. 2006;11:61-8.
13. Fowell SL, Southgate LJ, Bligh JG. Evaluating assessment: the missing link? Med Educ. 1999;33:276-81.

DOI: 10.5455/2320-6012.ijrms20150131

**Cite this article as:** Khan GN, Ishrat N, Khan AQ.

Using item analysis on essay types questions given in summative examination of medical college students: facility value, discrimination index. Int J Res Med Sci 2015;3:178-82.