

---

## Validation of Multiple True or False (MTF) Questions and Their Usefulness in Assessment in an Undergraduate Medical Programme

Vashe,A<sup>1</sup>., Ramnarayan,K<sup>2</sup>., Pallath,V<sup>3</sup>., Kamath,A<sup>4</sup>.

### Abstract

**Objective:** To find out difficulty index (P) and discrimination index (D) of the items - Multiple True or False (MTF) questions- set for the examinations conducted by constituent departments of a medical college, namely: anatomy, physiology, biochemistry, pathology, pharmacology and microbiology.

**Methods:** Scores obtained by students of first year MBBS (March 2010 batch) in anatomy, physiology, biochemistry and second year students (March 2009 batch) in pathology, pharmacology and microbiology were taken. MTF component of the examination of block 1, 2, 3 and 4 were considered. Difficulty index (P) and discrimination index (D) of MTF component of block 1, 2, 3 and 4 were analyzed using Microsoft excel. Correlations of MTF and essay scores of the examinations were also performed.

**Results:** The average easy questions( $P \geq 75\%$ ) were 48%, 65% and 51% & percentage of average difficult questions ( $P \leq 25\%$ ) were 3%, 1% and 2% in anatomy, physiology and biochemistry respectively. Anatomy, physiology and biochemistry had 71%, 56% and 67% of questions with discrimination index  $D \geq 0.2$ . The average easy questions were 64%, 51% and 46% & percentage of average difficult questions was 3%, 3% & 4% in pathology, pharmacology and microbiology respectively. Pathology, pharmacology and microbiology had 40%, 51% and 50% of questions with discrimination index  $D \geq 0.2$ . The essay and MTF scores of each subject showed strong and significant correlations.

**Conclusion:** The difficulty index obtained from the analysis shows that the questions were easy for the students. However, the study shows that about 50% of the questions were capable of discriminating students with higher ability from those with lower ability ( $D \geq 0.2$ ).

**Key words:** assessment, difficulty index, discrimination index

---

### Introduction

Assessment is an integral part of any education system. The main goal of conducting an assessment is to measure competency of students in the respective domains. One of the widely accepted ways of assessing students is by conducting written tests, where in, either objective or essay type questions will be asked and students will have to answer them within a pre-determined period of time. In such a method, quality of questions that are used to measure knowledge of students are very important in determining the success of assessment. A technique called item analysis is done to analyze the quality and utility of questions or items in differentiating the performance of students.

<sup>1</sup>Department of Physiology and <sup>3</sup>Department of Microbiology, Melaka Manipal Medical College (Manipal campus).

<sup>2</sup>Vice Chancellor, Manipal University

<sup>4</sup> Department of Community Medicine, Kasturba Medical College, Manipal, Karnataka, India.

#### Corresponding author:

<sup>3</sup>Dr.Asha. V,  
Dept. of Physiology, Melaka Manipal Medical College (Manipal Campus)  
Manipal University, Manipal- 576 104,  
Karnataka, India.

E-mail: [drashabhat@gmail.com](mailto:drashabhat@gmail.com)

Item analysis is a process to evaluate effectiveness of items in a test, based on

examinees responses to each item (Kim, 1997). Furthermore to assess effects of educational programs it is very important to conduct item analysis (Hetzl, 1997). It can also be used to obtain information such as difficulty and discriminatory power of items (Kim, 1999).

Difficulty Index (P) is the percentage of total correct responses to the test questions. Difficulty Index is analyzed to understand how easy or difficult the question is for the students (Singh et al., 2000). Higher the value of difficulty index, the lower is the difficulty of the question (Backoff et al., 2000; Hetzel, 1997). If the index is above 75% it is considered as an easy question while index below 25% is declared as a difficult question (Singh et al., 2000; Sim&Rasiah, 2006). For example, if the difficulty index of a question is 75%, it means 75% of students have answered the question correctly and hence it is considered as easy question. On the other hand, if the difficulty index of the question is 25%, it indicates that only 25% of students have answered the question correctly and obviously the question is considered difficult. Value of difficulty index ranges from zero to one (Zurawski, 1998). When no examinees answered correctly value of difficulty index is zero. It attains a value of one when all examinees answered correctly.

Discrimination Index (D): To find out the ability of a question to discriminate students based on their understanding of subject matter, discrimination index is calculated. Higher the discrimination index better is the question in discrimination. If a particular question is doing a good job of discriminating between those who score high and those who score low, more students in top scoring would have answered the question correctly (Hetzl, 1997). Discrimination index ranges from -1 to 1. An item everybody answered correctly or incorrectly will have zero discrimination index (Zurawski, 1998). If students in the lower group answer more, then, discrimination index will be negative. If more number of students in higher group answers correctly, then, discrimination index will assume a positive value.

Following are the guidelines regarding discrimination index mentioned by Ebel (1972).

- $D \geq 0.4$ ; very good
- $D: 0.3$  to  $0.39$ ; reasonably good, possibly subject to improvement
- $D: 0.2$  to  $0.29$ ; marginal, need some revision
- $D < 0.19$ ; poor, need major revision or to be eliminated.

Brown (1983) and Algina (1986) have found that test questions with discrimination index equal to or above 0.2 are acceptable and these questions differentiate upper and lower groups.

In recent years, MBBS curriculum at our institute has undergone many changes. In changed circumstances, it becomes necessary to explore the level of learning by students and their level of competency. An item analysis was carried out to determine the reliability of the tools of assessment which would enable us to meet the above two requirements.

Objective of the present study was to find out the difficulty Index and discrimination index of Multiple True or False (MTF) questions set for examinations conducted by constituent departments of our institution, namely: anatomy, physiology, biochemistry, pathology, pharmacology and microbiology.

## Methods

Our institution offers MBBS course which is of five years duration. Physiology, anatomy and biochemistry are taught in first year curriculum, while, pathology, pharmacology and microbiology are taught in the second year. The curriculum is taught in 4 blocks (teaching units). Block 1 includes basic concepts, skin, muscle, bones, joints and blood. Block 2 includes cardiovascular system, respiratory system, GIT, nutrition and hepatobiliary system. Block 3 includes endocrine, reproduction, kidney and electrolytes. Block 4 includes central nervous system, special senses and molecular biology. Duration of every block is 10 weeks. At the end of every block, examinations will be conducted which consists of both essay type (60 marks) and MTF (120 marks) type of questions. For the purpose of the present study, scores obtained by first year MBBS students (March 2010 batch) in anatomy, physiology, biochemistry and second year students (March 2009 batch) in pathology, pharmacology and microbiology was taken. Only MTF component of the examination of block 1, 2, 3 & 4 were considered and analysis was done using Microsoft excel. Two parameters, namely, Difficulty index (P) and Discriminatory index (D) were analyzed.

The formula used for evaluation of difficulty index:

$$P = R/T \text{ (Sim \& Rasiah, 2006)}$$

(Where, P was difficulty Index, R was total correct responses and T was total number of students appeared for the examination)

To analyze discrimination index, papers were arranged in rank order with students scoring highest marks positioned in the top. It was then divided into three equal groups with higher Ability Group (HAG), which was the top 1/3rd and Lower Ability Group (LAG), which formed the bottom 1/3rd of the group (Chandratilake *et al.*, 2010).

Discrimination index was calculated as

Discrimination index  $D = \frac{PHAG - PLAG}{T}$   
(Sim&Rasiah, 2006; Chandratilake *et al.*, 2010)

(Difficulty index of higher ability group – Difficulty index of lower ability group)

Correlation between MTF and essay scores was analyzed using Pearson correlation. This was performed to understand the reliability of the MTF scores in categorizing students as HAG and LAG in our settings.

**Results:** Table 1 shows the difficulty and discrimination index of questions in anatomy, physiology and biochemistry in different blocks. The average easy questions ( $P \geq 75\%$ ) were 48%, 65% and 51% & percentage of average difficult questions ( $P \leq 25\%$ ) were 3%, 1% & 2% in anatomy, physiology and biochemistry respectively. Anatomy, physiology and biochemistry had 71%, 56% and 67% of questions with discrimination index  $D \geq 0.2$  as shown in table 2.

As shown in the table 3, in second year the average easy questions were 64%, 51% and 46% & percentage of average difficult questions were 3%, 3% & 4% in pathology, pharmacology and microbiology respectively. Pathology, pharmacology and microbiology had 40%, 51% and 50% of questions with discrimination index  $D \geq 0.2$  as shown in table 4.

The tables 5 and 6 show correlation of students' scores in the essay-type questions and MTF questions. All the correlation coefficients (r) were above 0.6 in all subjects indicating a strong correlation and they were also significant ( $p < 0.001$ ).

**Table 1: Difficulty index of MTF questions asked for first year MBBS students (n=131)**

Department / Subject	Block	Difficulty Index (P) in Percentage	
		> 75 %	< 25 %
Anatomy	Block-1	46	4
	Block-2	45	1
	Block-3	54	1
	Block-4	47	4
Physiology	Block-1	64	0
	Block-2	59	0
	Block-3	74	0
	Block-4	62	3
Biochemistry	Block-1	43	5
	Block-2	50	1
	Block-3	62	2
	Block-4	48	0

**Table 2: Discrimination index of MTF questions asked for first year MBBS students (n=131)**

Department/ Subject	Block	Discrimination Index (D) in Percentage
		≥ 0.2
Anatomy	Block-1	64
	Block-2	74
	Block-3	69
	Block-4	75
Physiology	Block-1	52
	Block-2	66
	Block-3	48
	Block-4	59
Biochemistry	Block-1	74
	Block-2	69
	Block-3	63
	Block-4	62

**Table 3: Difficulty index of MTF questions asked for second year MBBS students**

Department / Subject	Block	Difficulty	Index (P) in
		> 75 %	< 25 %
Pathology	Block-1	56	3
	Block-2	55	1
	Block-3	63	3
	Block-4	83	3

Pharmacology	Block-1	44	5
	Block-2	45	1
	Block-3	52	2
	Block-4	63	2
Microbiology	Block-1	47	5
	Block-2	40	3
	Block-3	39	3
	Block-4	58	1

**Table 4: Discrimination index of MTF questions asked for second year MBBS students**

Department/ Subject	Block	Discrimination Index (D) in Percentage	
		≥ 0.2	
Pathology	Block-1	45	
	Block-2	44	
	Block-3	40	
	Block-4	32	
Pharmacology	Block-1	54	
	Block-2	51	
	Block-3	51	
	Block-4	47	
Microbiology	Block-1	51	
	Block-2	54	
	Block-3	53	
	Block-4	42	

**Table: 5 Correlation of essay and MTF scores of block theory examinations of first year MBBS students**

	Anatomy		Physiology		Biochemistry	
	Correlation coefficient ( <i>r</i> )	p	Correlation coefficient ( <i>r</i> )	P	Correlation coefficient ( <i>r</i> )	p
Block 1	0.821	< 0.001	0.815	< 0.001	0.828	< 0.001
Block 2	0.776	< 0.001	0.825	< 0.001	0.868	< 0.001
Block 3	0.757	< 0.001	0.786	< 0.001	0.746	< 0.001
Block 4	0.750	< 0.001	0.824	< 0.001	0.792	< 0.001

**Table: 6 Correlation of essay and MTF scores of block theory examinations of second year MBBS students**

	Pathology		Pharmacology		Microbiology	
	Correlation coefficient ( <i>r</i> )	p	Correlation coefficient ( <i>r</i> )	P	Correlation coefficient ( <i>r</i> )	P
Block 1	0.722	< 0.001	0.771	< 0.001	0.783	< 0.001
Block 2	0.784	< 0.001	0.752	< 0.001	0.705	< 0.001
Block 3	0.695	< 0.001	0.662	< 0.001	0.704	< 0.001
Block 4	0.680	< 0.001	0.743	< 0.001	0.746	< 0.001

### Discussion

The study shows a consistent level of difficulty and discrimination indices being maintained from subject to subject. The students' score in MTF component correlates strongly with the essay component in all subjects throughout the year. Such consistency can be attributed to the teaching-learning and assessment process followed in our medical school. The question papers set by the faculty members for examinations were scrutinized at multiple levels. Question papers were reviewed by a team of experts involved in that particular block which were then thoroughly scrutinized by head

of the department. This ensures that the questions were confined to the objectives defined for the course and also maintains the standard of question paper.

It can be seen from the analysis that physiology from first year and pathology from second year had more easy questions compared to other subjects. The subject content and its closeness to the day to day life might have had a strong influence on the outcome. Among that the block which deals with endocrinology, reproductive system and kidney- third block had more easy questions. In this block in physiology, students have to collect pictures about the endocrine

disorders and paste it in the journal book; also they have to solve renal problems. This might have made the students to learn more and the reason for more students answering and so the lower discrimination.

The study of a preclinical question paper done by Mitra *et al.*, (2009) found that 40% of the total test questions had difficulty index crossing 80% and another study of year two examinations of a medical school reported by Si MuiSimet *et al.*, (2006), also found that about 40% of the MCQ items crossed difficulty index 75%. Compared to these studies, we have more easy questions. If the question is easy for the students it also means that they have learnt the topic (Zurawski, 1998). It is substantiated by the good correlation of students' essay and MTF scores. Thus, higher competency level of the learners may be one reason for the outcome. Teaching methodology adopted at our institution involved active learning components such as Self Directed Learning (SDL), Problem Based Learning (PBL) in addition to didactic lectures. This forced students to do additional self-learning which invariably involved referring various learning materials related to the subject. This might have improved competency, knowledge and retention level of students (Dolmans & Schmidt, 1996).

Our institution provided very good academic support for students where in, weak students were identified by faculty members and were given proper directions to improve their competency level. Additional efforts were put by faculty members to see that students learn as per the objectives set in each subject. Printed materials clearly defining the objectives set in each subject were issued to students well in advance, to keep them informed of what was expected to be learnt as per the curriculum. This might have helped students to prepare themselves better to meet expectations and obtain a wide understanding of subjects, which in turn, increased probability of answering questions correctly. These could be the other reasons for large percentage of students answering the questions correctly. Carroll (1993) quotes that "The item difficulty and discrimination are often reciprocally related". This may be the reason for the low discrimination index.

## Conclusion

The difficulty index obtained from the analysis shows that the questions were easy for the students. However, the discriminatory index decreases as the difficulty index increases, but the study shows that about 50% of the

questions were capable of discriminating students with higher ability from those with lower ability ( $D \geq 0.2$ ). If students' success in examination is to be considered as an indicator of effective teaching learning process, then the assessment should be a reliable indicator of student achievement of intended learning outcomes of the course.

## Acknowledgment

We acknowledge the dean and associate dean of MMMC, Manipal, Head of the dept. of anatomy, physiology, biochemistry, pathology, pharmacology and microbiology for giving permission to conduct this study.

## References

- Backhoff, E., Larrazolo, N. & Rosas, M. (2000) The level of difficulty and discrimination power of the Basic Knowledge and skills Examination (EXHCOBA) [Online]. *Revistaelectronca de investigation Educativa* Available at: <http://redie.uabc.mx/vol2no1/contents-backhoff.html> [Accessed December 21, 2010].
- Brown, F.G. (1983) *Principles of educational and psychological testing* (3<sup>rd</sup>ed.) New York: Holt, Rinehart and Winston.
- Carroll, R.R. (1993) Evaluation of vignette-type examination items for testing medical physiology, *American Journal of Physiology*, 264, pp. S11-5.
- Chandratilake, M.N., Davis, M.H. & Ponnampereuma, G. (2010) Evaluating and designing assessments for medical education: the utility formula, *The Internet Journal of Medical Education* [Online]. Available [http://www.ispub.com/journal/the\\_internet\\_journal\\_of\\_medical\\_education/volume\\_1\\_number\\_1\\_74/article/evaluating\\_and\\_designing](http://www.ispub.com/journal/the_internet_journal_of_medical_education/volume_1_number_1_74/article/evaluating_and_designing) [Accessed May 8, 2011]
- Crocker, L. & Algina, J. (1986) *Introduction to classical and modern test theory*. New York: Holt, Rinehart and Winston.
- Dolmans, D. & Schmidt, H. (1996) The advantages of problem-based curricula, *Postgrad Medical Journal*, 72(851), pp. 535-538.
- Ebel, R.L. (1972) *Essentials of educational measurement* (1st ed.) New Jersey: Prentice Hall.
- Hetzel, S.M. (1997) Basic concepts in item and test analysis [Online]. Available at <http://ericae.net/ft/tamu/Epsy.htm> (Accessed May 5, 2011).
- Kim, S.H. (1999) A computer program for classical item analysis [Online]. Available at <http://shkim.myweb.uga.edu/cia.htm> [Accessed May 5, 2011].

Mitra, N.K., Nagaraja, H. S., Ponnudurai, G. & Judson, J.P (2009) The levels of difficulty and discrimination indices in type A multiple choice questions of pre-clinical semester 1 multidisciplinary summative tests, *International e-Journal of Science, Medicine & Education*, 3 (1), pp. 2 - 7.

Sim, S.M. & Rasiah, R.I. (2006), Relationship between item difficulty and discrimination indices in true/false-type multiple choice questions of a para clinical multidisciplinary paper, *Ann Academy of Medicine Singapore*, 35, pp. 67-71.

Singh, T., Singh, D. & Vinod, K.P. (2000) Test and item Analysis. *Principles of Medical Education* (1<sup>st</sup>ed.) IAP Education Centre, Dept. Of Pediatrics, AIIMS, New Delhi.

Zurawski, R.M. (1998) Making the most of exams: Procedures for item analysis[Online]. Available at [http://www.ntlf.com/html/pi/9811/exams\\_1.htm](http://www.ntlf.com/html/pi/9811/exams_1.htm) (Accessed May 5, 2011)