International Journal of Anatomy and Research,
Int J Anat Res 2015, Vol 3(4):1607-10. ISSN 2321- 4287
DOI: http://dx.doi.org/10.16965/ijar.2015.299

Original Research Article

RELATIONSHIP BETWEEN DIFFICULTY INDEX AND DISTRACTER EFFECTIVENESS IN SINGLE BEST-ANSWER STEM TYPE MULTIPLE CHOICE QUESTIONS

Pradip Chauhan *[1], Girish Rameshbhai Chauhan[2], Bhoomika Rameshbhai Chauhan[3], Jayesh Vrajlal Vaza [4], Suresh Punjabhai Rathod [5].

[*1] Department of anatomy, Pandit Dindayal Upadhyay Government Medical College, Gujarat, India.
[2] Assistant Professor, Department of Oral Pathology, Government Dental College, Gujarat, India.
[3] Assistant Professor, Department of Gynaecology, GMERS Medical College, Sola, Gujarat, India.
[4] Assistant professor, Department of Orthopaedics, L.G. Medical college, Gujarat, India.
[5] Professor and Head, Department of Anatomy, P.D.U. Medical College, Gujarat, India.

ABSTRACT

Background: Single best-answer multiple-choice questions (MCQs) consist of a question (the stem) two or more choices from which examinees must choose the correct option (the distracters) and one correct or best response (the key). Item analysis is the process of collecting, summarizing and using information from students' responses to assess the quality of test items. Classical test theory for item analysis is most followed method to determine the reliability by calculating Difficulty Index (P score) and Discriminating Index (D score) and Distracter effectiveness

Aim: This Study was aimed to calculate P scoreand distracter effectiveness; to find out relationship between P score and distracter effectiveness.

Material and methods: In this Cross Sectional study 65 items responded by 120 Students of first year M.B.B.S were studied for Item Analysis. Difficulty Index, and Distracter Effectiveness were calculated for each item. Distracters were identified and classified as Functioning and Non- functioning distracter. Interrelationship between P Score, and Distracter Effectiveness was calculated and analyzed by Epinifo 7 software

Result: We found Items with two functioning distracters were more difficult than that of others followed by items with three functioning distracters.

Conclusion: Distractors affect the item difficulty index and by the means also affects quality of the assessment.

Key words: Item Analysis, Difficulty Index, Distracter Effectiveness, Multiple Choice Question.

Address for Correspondence: Dr. Pradip Rameshbhai Chauhan, Department of anatomy, Pandit Dindayal Upadhyay Government Medical College, Rajkot, Gujarat, India. Mob. No.: +918866199560
E-Mail: prajjawalitresearch@gmail.com

Access this Article online

Quick Response code

DOI: 10.16965/ijar.2015.299

Web site: International Journal of Anatomy and Research
ISSN 2321-4287
www.ijmhr.org/ijar.htm

Received: 14 Oct 2015 Accepted: 16 Nov 2015
Peer Review: 14 Oct 2015 Published (O): 30 Nov 2015
Revised: None Published (P): 31 Dec 2015

INTRODUCTION

Multiple-choice questions (MCQs) are used more and more in departmental examinations or as comprehensive Examinations at the end of an academic session to determine progress or to make decisions regarding the certification

of a candidate [1-4]. Single best-answer multiple-choice questions (MCQs) consist of a question -the stem, two or more choices from which examinees must choose the correct option-the distracters, and one correct or best response-the key [5]. Item analysis is the process of collecting, summarizing and using information from students' responses to assess the quality of test items [3,6]. Item analysis is a valuable, yet relatively simple, procedure performed after the examination that provides information regarding the reliability and validity of a test item [7].

Classical test theory is the most followed method to determine the reliability by calculating Difficulty Index {P score} and Discriminating Index {D score} [6]. Difficulty index also called ease index, describes the percentage of students who correctly answered the item. It ranges from 0 - 100%; the higher the percentage, the easier the item. The recommended range of difficulty is from 30-70%. Items having p-values below 30% and above 70% are considered difficult and easy items respectively [8].

Discrimination index (DI), also called point biserial correlation (PBS), describes the ability of an item to distinguish between high and low scorers [9]. It ranges between -1.00 and +1.00; discriminating index $\geq$ 0.35 is considered as excellent discriminator, DI $\leq$ 0.20 is poor discriminator. Discriminating Index is < 0.00, indicates low performing students got a specific item correct more often than the high scorers [7]. Analysis of distracters is also important technique that provides information regarding the individual distracters and the key of a test item. Using these tools, the examiner is able to modify or remove specific items from subsequent exams [10].

Aims and Objectives: This study was conducted with aim to find out relationship between difficulty index, distracter effectiveness. For the aim Difficulty Index of each item was calculated; distracter effectiveness for each option was calculated. Frequency of functioning and non-functioning distracters was assessed.

MATERIALS AND METHODS

Data collection: In this cross sectional study, test with 0.87 (>0.70) KR20 reliability index was selected; which was taken in four different session during year 2012-13 for anatomy subject in P.D.U. Medical College, Rajkot. Three of four sessions contained 15 items and one session contained 20 items. Items are defined as Stem type single best choice of four options Multiple Choice Questions, constructed confidentially by experienced senior teachers of department and scrutinized by examination committee of the Anatomy Department. There was not any penalty or negative marking for wrong response. 120 students who attended all sessions were included in study. Item analysis: Item Difficulty index refers to the percentage of the total number of correct responses to the test item-calculated by following formula.

Difficulty index (P score) = (Total true response of the item / Total responses) ×100

Where Total responses= (True response + Wrong response + Not responded)

Discriminating index: Total score by sum of all four sessions for each individual student was calculated and arranged in descending order from highest score to lowest score. Upper 1/3 students (Highest scorer 40 students out of 120) were selected to include in higher group (H) and lower 1/3 students (lowest scorer 40 students out of 120) were selected to include in lower group (L). Item Discriminating index is calculated by following formula

Discriminating index (D Score)

= (HT-LT / T) ×2

Where HT = Number of correct responses in Upper Group, LT = Number of correct responses in Lower Group and T= Total number of responses in both group.

Similarly discriminating power for each distracter was calculated by following formula

Discriminating power for distracter

= (HR-LR / T) ×2

Where HR = Number of responses from Upper Group, LR = Number of responses from Lower Group and T= Total number of responses

Distracter effectiveness for the option indicates percentage of students choosing the option of item as an answer. Good distracters appeal to a higher proportion of low achieving examinees when compared with high-achieving examinees,

thereby resulting in a negative statistic. The advantage of the discrimination index is that it is simple to compute and explain. Therefore, a non-functioning distracter was defined as an option with either a response frequency of <5% or a positive discriminating power.

Data analysis: Frequency distribution was done for number of functioning distracter for each item. Relationship between difficulty index and number of functioning distracter was calculated. Data was calculated and analyzed by Microsoft excel 2007 and Epi info version 7 software.

RESULTS AND DISCUSSION

In this study 65 items responded by 120 students were analyzed. Mean Item Difficulty Index was 57.77 (43.49-72.05). (p value<0.05, confidence interval >95%).

95% (247 out of 260) distracters have effectiveness was ≥ 5%. 63.85% (166 out of 260) distracters were functional (Table: 1a, 1b).

Items with two functioning distracters were more difficult than that of others followed by items with three functioning distracters (Table: 2). Mean Difficulty Index of Items with functioning distracters two and three was in ideal range.

Mean difficulty Index of test items analyzed in the study was 57.77 (43.49-72.05) which is interpreted as ideal [8]. This Study results show better quality of Items than the study of Si Mui Sim et al (2006), who found that about 40% of the MCQ items crossed difficulty index 70% showing that the test items were easy for the examinees [11]. Mitra N K et al (2009), found mean difficulty index scores of the individual tests for his study ranging from 64% to 89% [12].

Mean number of functional distracters (2.55) in the study was higher than study of Tarrant M et al (2009) and Shah CJ (2011) [10,13]. The study results show two or three functional distracters are optimal to format quality MCQs; items with four functioning distracters couldn't be analyzed because of inappropriate number. Items with two or three options were ideally difficult; results support the previous studies of Trevisan et al. (1991) Who found that three-option items were more discriminating and had fewer items with non-performing distracters than five-option items [14].

Tarrant M (2009) concluded items with two functioning distracters were more difficult than items with three functioning distracters [10]. In this study difficulty Index for items with two and three functioning distracters was not significantly different.

Only one item was with four functioning distracter in the study. Not a single item had 0 number of functioning distracter. There was no penalty or negative marking in the test.

Table 1a: Distracter frequency distribution as per Distracter effectiveness and Distracter Difficulty Index.

Number of items= 65		Number of distracters = 260	
Distracter effectiveness (frequency):	n (%)	Distracter discrimination power	n (%)
≥5%	247 (95%)	≥0	85 (32.70%)
<5%	12 (4.62%	<0	175(67.30%)
=0%	1 (0.38%)		

Table 1b: Frequency distribution according to number of functioning distracter.

Functioning distracters	166 (63.85%)
Non functioning distracters	94 (36.15%)
Functioning distracters per item	n (%)
1	5 (7.69%)
2	20 (30.77%)
3	39 (60.00%)
4	1 (1.54%)
Mean (Functioning Distracters per item)	2.55 (1.23-3.87)

Table 2: Comparison of other studies for site of origin of profunda femoris artery.

Functioning distracters	Observations (n=65)	Item Difficulty Index Mean (± SD)
1	5	80.25 (70.36-90.14)
2	20	53.5 (37.79-69.21)
3	39	57.08 (45.87-68.29)
4	1	60
Total	65	57.77 (43.49-72.05)

(ANOVA Test)P value=0.00038 < (0.05),
Items with four functioning distracters were not included in ANOVA test
Mean difficulty Index is significantly different for number of functioning distracters per item

Limitations:

Study conducted in the medical institute may reflect the specified stream (medical) institute and cannot be applied for academic institutes of all streams.

CONCLUSION

Well-chosen distracters improve the quality of MCQs by affecting Item Difficulty Index. Regular assessment of all exams should be done to maintain quality of examination system. Further study is necessary to identify effect various numbers of distracters on quality of MCQs and test papers.

ACKNOWLEDGEMENTS

Authors are thankful to authority of P.D.U Government Medical College, Rajkot for giving permission and providing infrastructure to carry out this research project.

Conflicts of Interests: None

REFERENCES

[1]. Hubbard JP, Clemans WV. Multiple-choice Examinations in Medicine. A Guide for Examiner and Examinee. Philadelphia: London. Lea & Fabiger; 1961:180.

[2]. Skakun EN, Nanson EM, Kling S, Taylor WC. A preliminary investigation of three types of multiple choice questions. Med Edu 1979;13: 91-96.

[3]. Kehoe J. Basic item analysis for multiple choice tests. Practical Assessment, research, Evaluation 1995;4: 10. URL: http://PAREonline.net/getvn.asp?v=4&n=10 [accessed on 17 March 2014]

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

[5]. Cizek GJ, O'Day DM. Further investigations of nonfunctioning options in multiple-choice test items. Educ Psychol Meas 1994; 54(4):861-872. URL: http://www.biomedcentral.com/1472-6920/9/40 [accessed on 17th March,2014]

[6]. Zubairi AM, Kassim NLA.Classical and Rasch analysis of dichotomously scored reading comprehension test items. Malaysian Journal of ELT Research 2006;2: 1-20. URL: http://www.melta.org.my/modules/tinycontent/Dos/AinolMadziah.pdf [accessed on 17 March 2014]

[7]. 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. URL: http://dro.deakin.edu.au/eserv/DU:30003091/considine-designformatvalidity-2005.pdf [accessed on 17th march,2014]

[8]. Miller MD, Linn RL, Gronlund NE. Measurement and assessment in teaching. 10th edition. Upper Saddle River, NJ: Prentice Hall; 2009:576.

[9]. Fowell SL, Southgate LJ, Bligh JG. Evaluating assessment: the missing link? Med Educ 1999;33: 276-81.

[10]. Tarrant M, Ware J, Mohammed AM. An assessment of functioning and non-functioning distracters in multiple-choice questions: a descriptive analysis. BMC Med Educ 2009;9: 40. URL: http://www.biomedcentral.com/1472-6920/9/40 [accessed on 17th March.2014]

[11]. Si-Mui Sim, 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. URL: http://www.annals.edu.sg/pdf/35VolNo2200603/V35N2p67.pdf [accessed on 17th March.2014]

[12]. Mitra NK, Nagaraja HS, Ponnudurai G, Judson JP. The Levels of Difficulty and Discrimination Indices In Type A Multiple Choice Questions Of Pre-clinical Semester 1 Multidisciplinary Summative Tests. IEeJSME 2009;3 (1): 2-7. URL: http://web.imu.edu.my/ejournal/approved/iMEC_2.original_p02-07.pdf [accessed on 17th March.2014]

[13]. Shah CJ, Baxi SN, Parmar RD, Parmar D, Tripathi CB. Item Analysis of MCQ from Presently Available MCQ Books. Indian Journal For The Practising Doctor 2011;6(1);26-30

[14]. Trevisan MS, Sax G, Michael WB. The effects of the number of options per item and student ability on test validity and reliability. Educ Psychol Meas 1991;51(4):829-837. URL: http://www.psychologie-aktuell.com/fileadmin/download/ptam/2-2011_20110622/02_Baghaei.pdf [accessed on 17th March.2014]

How to cite this article:

Pradip Chauhan, Girish Rameshbhai Chauhan, Bhoomika Rameshbhai Chauhan, Jayesh Vrajlal Vaza, Suresh Punjabhai Rathod. RELATIONSHIP BETWEEN DIFFICULTY INDEX AND DISTRACTER EFFECTIVENESS IN SINGLE BEST-ANSWER STEM TYPE MULTIPLE CHOICE QUESTIONS. Int J Anat Res 2015;3(4):1607-1610. **DOI:** 10.16965/ijar.2015.299