International Journal of Anatomy and Research,
Int J Anat Res 2018, Vol 6(2.1):5156-62. ISSN 2321-4287
DOI: https://dx.doi.org/10.16965/ijar.2018.143

Original Research Article

MULTIPLE CHOICE QUESTIONS - ROLE IN ASSESSMENT OF COMPETENCY OF KNOWLEDGE IN ANATOMY

Padamjeet Panchal [1], Bheem Prasad *[2], Sarita Kumari [3].

[1,*2] Assistant Professor, Department of Anatomy, All India Institute of Medical Sciences, Phulwari Sharif, Patna, Bihar, India.

[3] Senior Resident, Department of Anatomy, All India Institute of Medical Sciences, Phulwari Sharif, Patna, Bihar, India.

ABSTRACT

Background: The accurate, reliable and timely assessment of students is an essential domain of teaching during Medical professional courses. The Multiple Choice Questions (MCQ) are time tested method of ready assessment of undergraduate students. Although it evaluates student's cognitive knowledge but does not evaluate professional skills. However it is said that MCQs emphasize recall of factual information rather than conceptual understanding and interpretation of concepts.

Objectives: The main objective of the study is to analyse the items with the help of item analysis and select the items which are good for incorporation into future question bank with reliability.

Materials and Methods: This study was done in Department of Anatomy, AIIMS, Patna. A 396 first year MBBS students of different batches took the MCQ test comprising 60 questions in two sessions. During the evaluation process of MCQ's each correct response was awarded one mark and no marks was awarded for any incorrect response. Each item was analysed for difficulty index, discrimination index and distractor effectiveness.

Results: The overall mean of Facilitative value, Discrimination Index, Distractor Effectiveness and Correlation Coefficient was 66.09 (±21.55), 0.26 (±0.16), 18.84 (±10.45) and 0.55±0.22 respectively.

Conclusion: The framing of MCQ should be according to Bloom's classification to assess cognitive, affective as well as psychomotor domain of the students. The MCQ having poor and negative discrimination should be reframed and again should be analysed.

KEY WORDS: Item Analysis, Multiple Choice Question, Difficulty Index, Discrimination Index, Distractor Effectiveness, Correlation Coefficients.

Address for Correspondence: Dr. Bheem Prasad, Assistant Professor, Department of Anatomy, All India Institute of Medical Sciences, Phulwari Sharif, Patna, Bihar-801507, India.
E-Mail: drbheemp@aiimspatna.org

Access this Article online	Journal Information
Quick Response code	**International Journal of Anatomy and Research**
	ICV for 2016 90.30
	ISSN (E) 2321-4287 | ISSN (P) 2321-8967
	https://www.ijmhr.org/ijar.htm
	DOI-Prefix: https://dx.doi.org/10.16965/ijar
DOI: 10.16965/ijar.2018.143	**Article Information**
	Received: 13 Feb 2018 Accepted: 07 Mar 2018
	Peer Review: 14 Feb 2018 Published (O): 05 Apr 2018
	Revised: None Published (P): 05 Apr 2018

INTRODUCTION

The accurate, reliable and timely assessment of students is an essential domain of teaching during medical professional courses. The Multiple Choice Questions (MCQ) are time tested method of ready assessment of undergraduate students. Although it evaluates student's cognitive knowledge but does not evaluate professional skills. However it is said that MCQs emphasize recall of factual information rather than conceptual understanding and interpretation of concepts [1].

Properly constructed MCQs can assess higher cognitive processing of Bloom's taxonomy such

as interpretation, synthesis and application of knowledge, instead of just testing recall of isolated facts [2, 3]. Test with MCQ and analysing their options have become the choice of many examiners in medical colleges [4]. Haladyna et al. reviewed the validity of taxonomy of MCQ tests and wrote the guidelines for them [5]. Gajjar et al. examined the quality of MCQ tests and emphasized that a good MCQ truly assess the knowledge and was able to differentiate the students of different abilities [6]. While Sharif et al. concluded that MCQ was an efficient tool for measuring the achievement of learners [7]. Even Vyas and Supe suggested that MCQs with three alternatives should be preferred than the four or five options [8].

There are three components of a MCQ, direction (instruction to the students), stem (the question) and choices (alternatives). The correct alternative is called as answer and the other alternatives are called as distracters [9]. To assess the different domains it is important to have a good items. Item analysis is a process which assesses the quality of those items and of the test as a whole [10]. A well-constructed MCQ can differentiate between low achievers and high achievers. The main objective of the study is to analyse the items with the help of item analysis and select the items which are good for incorporation into future question bank with reliability. These items are able to discriminate good and poor performing students. So it will be helpful to improve our teaching learning process and the low achievers will be given more imperative teaching. A valid question bank is a must for proper assessment. The ultimate aim of any teaching will remain futile unless and until it is assessed. By MCQ there is analysis of the teaching learning process and by item analysis the analysis of the assessment method is done. The present study was done with a motto to analyse the role of MCQ in testing the competency of knowledge in anatomy.

MATERIALS AND METHODS

This study was done in Department of Anatomy, All India Institute of Medical Sciences, Phulwari Sharif, Patna, Bihar. The study includes the analysis of the Anatomy MCQ's attempted by the student of 1st year MBBS students of year 2013-2017. The MCQs were constructed by subject experts and they were given as a part of (section A) of whole examination. The MCQ's response sheet attempted by the student during the first year professional examination were taken year wise from the Examination section, AIIMS, Patna for MCQ's analysis after taking approval from the ethics committee of the Institution. The MCQ's analysed were of single best response type having four alternatives. There were three distractors and one correct response known as key. The MCQ were taken into two separate sessions as SA1 (Section A: SA1) and SA2 (Section A: SA2) respectively. SA1 contained MCQ's based on topics from general anatomy, general histology, general embryology, upper limb, thorax head and neck region while SA2 contained MCQ's based on topics from systemic histology, systemic embryology, genetics, abdomen, pelvis and lower limb. In each SA 30 MCQ's were analysed with 90 distractors and 30 keys. During the evaluation process of MCQ's each correct response was awarded one mark and no marks was awarded for any incorrect response. The data obtained was analysed in excel worksheet of Microsoft office ver. 2013. Each year section wise, depending upon the scores achieved by the students, they were arranged in merit list and categorised into two groups. The higher ability (H) group consisting of students of upper one third part of merit list, lower ability (L) group consisting of students of lower one third part of merit list and the middle 1/3 part of the merit list of students were excluded from the study.

Difficulty index (FV): The FV decides the difficulty of items. If the FV is <30, it indicates that item is difficult, between 30-70 is acceptable range and >70 is easy. Each MCQ was analysed with formula, FV: H+L/N x100

Where, H= No. of Student answering correctly in high achiever's group, L= No. of students answering correctly in low achiever's group and N= Total No. of students.

Discrimination index (DI): Discriminates between high and low achievers. If it is <0.20 that means the item is having poor discriminative power, between 0.2 -0.4 means good discriminative power. >0.4 means having excellent discriminative power. Negative DI means the

low achievers have marked more correct response than the high achievers. Discrimination index was analysed by following formula.

DI: (H-L/N) x2

Distractor effectiveness (DE): DE means distractor effectiveness that means the alternatives other than key are how much distracting the students. Those who distract more than 5% students are considered to be functional. Any item with one functional distractor is considered to be 33% effective, with two functional distractor 66% effective and with 3 functional distractor 100% effective.

DE: (H+L/N) x100

Correlation coefficient (Kr20**):** It is the coefficient which measures reliability of items with binary variables reliability refers to the amount of consistency of the results of a test. The score varies from 0 to 1. 0 is no reliability and+1 is perfect reliability. Kr20 score closer to 1, the more reliable the test. Score above 0.5 is considered reasonable,

Kr20=(k/k-1)x(1-Σ q)/var)

Where: k=sample size for the test, var= variance for the test, p=proportion of people passing the item, q=proportion of people failing the item,

Σ= sum up

Inference- FV was categorised into three parts. <30 is considered to be difficult, 30-70 considered to be acceptable and >70 is easy.

RESULTS

Figure 1 showed that out of all the papers of section 1 (SA1) 2014 and out of all the papers of section 2 (SA2) 2017 was having maximum numbers of acceptable items.

Figure 2 shows the discriminative capacity of items between high and low scorers. Out of all the papers of section 1 (SA1) 2015 and out of all the papers of section 2 (SA2) 2013 was having maximum numbers of items with good to excellent discrimination index.

Figure 3 shows the numbers of functional distractors. Out of all the papers of section 1 (SA1) 2017 and out of all the papers of section 2 (SA2) 2016 was best on the basis of maximum numbers of functional distractors.

Table 1 shows correlation between FV and DE.

The DE distribution among acceptable range of FV, were seen best in SA1 and SA2 of year 2017 because they are having maximum no. of items with 3 functional distractor.

Table 2 shows the correlation between FV and DI. The DI among the acceptable and easy range of FV, were seen comparatively better in SA1 of years 2014 and 2015. Whereas in case of SA2 papers, it is best seen in year 2017 because they are having maximum no. of items with good discrimination power.

Table 3 shows mean (± SD) of difficulty index, discrimination index and distractor effectiveness. On the basis of mean values we found that out of all the papers of section 1 (SA1) 2017 and out of all the papers of section 2 (SA2) 2016 was the best according to FV. According to discrimination index SA1 of 2015 and SA2 of 2013 were the best and according to distractor effectiveness SA1 of 2017 and SA2 of 2016 were best.

Figure 4 shows that section 1 (SA1) of 2015 and section 2 (SA2) of 2014 were best as per the values of correlation coefficient.

Fig. 1: Distribution of Items in relation to FV.

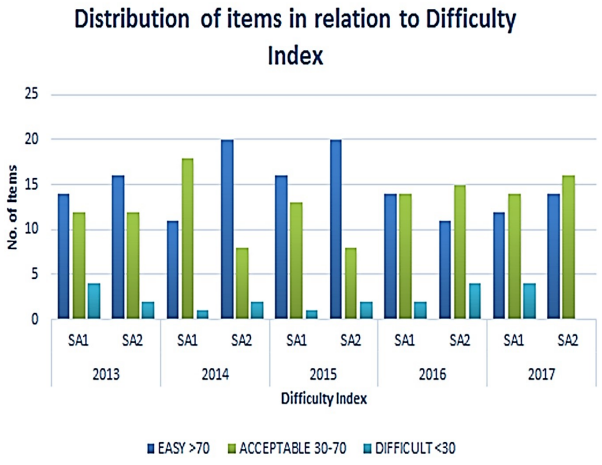

Fig. 2: Distribution of Items in relation to discrimination index.

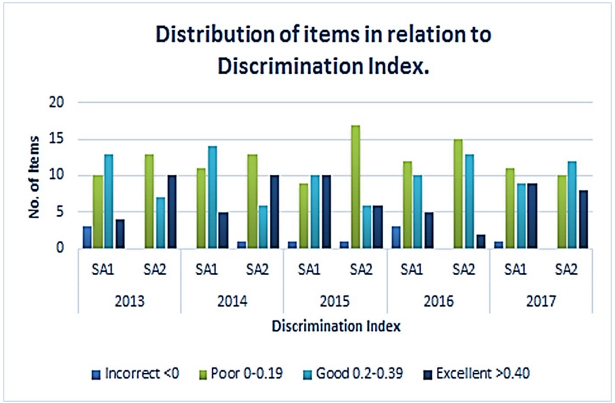

Fig. 3: Distribution of year wise functional distractors of SA1 and SA2.

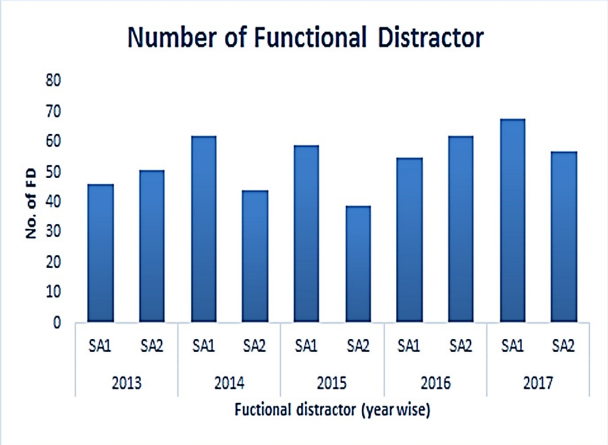

Fig. 4: Distribution of year wise correlation coefficient of SA1 and SA2.

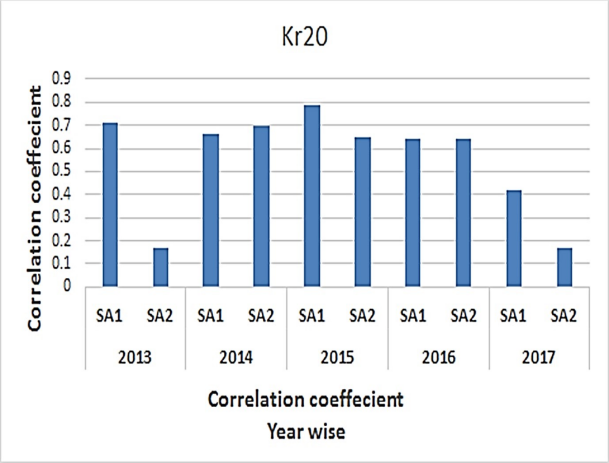

Table 1: Correlation between FV and DE.

Year	SECTION	FV	FD				Total
			3FD	2FD	1FD	0FD	
2013	SA1	<30	1	0	3	0	4
		30-70	5	2	5	0	12
		>70	0	4	8	2	14
		Total	6	6	16	2	30
	SA2	<30	0	2	0	0	2
		30-70	1	10	1	0	12
		>70	2	5	7	2	16
		Total	3	17	8	2	30
2014	SA1	<30	1	0	0	0	1
		30-70	5	10	2	1	18
		>70	5	1	4	1	11
		Total	11	11	6	2	30
	SA2	<30	1	0	1	0	2
		30-70	5	2	1	0	8
		>70	0	7	6	7	20
		Total	6	9	8	7	30
2015	SA1	<30	1	0	0	0	1
		30-70	9	4	1	0	14
		>70	2	4	6	3	15
		Total	12	8	7	3	30
	SA2	<30	0	1	1	0	2
		30-70	5	3	0	0	8
		>70	2	1	7	10	20
		Total	7	5	8	10	30
2016	SA1	<30	1	1	0	0	2
		30-70	9	5	0	0	14
		>70	0	2	9	3	14
		Total	10	8	9	3	30
	SA2	<30	2	2	0	0	4
		30-70	8	5	2	0	15
		>70	0	6	4	1	11
		Total	10	13	6	1	30
2017	SA1	<30	4	0	0	0	4
		30-70	9	5	0	0	14
		>70	2	5	3	2	12
		Total	15	10	3	2	30
	SA2	<30	0	0	0	0	0
		30-70	10	5	1	0	16
		>70	2	3	4	5	14
		Total	12	8	5	5	30

Table 2: Correlation between FV and DI.

Year	SECTION	FV	DI				Total
			<0	0-0.19	0.2-0.39	>0.4	
2013	SA1	<30	2	2	0	0	4
		30-70	0	3	6	3	12
		>70	1	5	7	1	14
	SA2	<30	0	2	0	0	2
		30-70	0	2	4	6	12
		>70	1	9	3	4	16
2014	SA1	<30	0	1	0	0	1
		30-70	0	5	9	4	18
		>70	0	5	5	1	11
	SA2	<30	1	1	0	0	2
		30-70	0	1	1	6	8
		>70	0	11	5	4	20
2015	SA1	<30	0	0	0	1	1
		30-70	0	2	4	7	13
		>70	1	7	6	2	16
	SA2	<30	0	2	0	0	2
		30-70	0	1	3	4	8
		>70	1	14	4	2	20
2016	SA1	<30	1	1	0	0	2
		30-70	0	4	5	5	14
		>70	2	7	5	0	14
	SA2	<30	0	4	0	0	4
		30-70	0	6	7	2	15
		>70	0	6	5	0	11
2017	SA1	<30	0	0	3	1	4
		30-70	1	4	4	5	14
		>70	0	7	2	3	12
	SA2	<30	0	0	0	0	0
		30-70	0	7	5	4	16
		>70	0	3	7	4	14

Table 3: Range, mean, Standard deviation and Standard error of Difficulty index, Discrimination Index and Distractor Effectiveness.

Year	Section	FV		DI		DE	
		Range	Mean± S.D.	Range	Mean± S.D.	Range	Mean± S.D.
2013	SA1	4.17 -91.67	63.89±24.38	-0.23	0.22±0.15	0-91.67	12.08±17.44
	SA2	8.33 -100	64.44±21.43	0-0.21	0.41±0.21	0-1.08	0.08±0.21
2014	SA1	17.65 -91.18	63.24±18.73	0-0.15	0.26±0.15	0-60.29	12.38±11.96
	SA2	13.24 -98.53	74.46±21.83	0.03-0.2	0.25±0.2	0-0.81	0.09±0.12
2015	SA1	27.27 -97.73	68.11±19.51	0.05-0.17	0.29±0.17	0-36.36	10.42±9.44
	SA2	6.82 -100	74.09±25.09	0.05-0.18	0.2±0.18	0-79.55	8.64±13.02
2016	SA1	29.69 -98.44	67.34±22.43	0.09-0.17	0.22±0.17	0-67.19	11.38±13.3
	SA2	9.38 -92.19	58.44±22.25	0-0.13	0.23±0.13	0-76.56	14.26±14.19
2017	SA1	3.13 -98.44	57.55±24.06	0.13-0.19	0.26±0.19	0-51.56	14.2±12.41
	SA2	40.63 -95.31	69.38±15.86	0.09-0.13	0.29±0.13	0-93.75	10.71±12.49

DISCUSSION

The present study was undertaken to analyse the MCQs with the help of item analysis and select the items which are good enough to assess the knowledge of subject anatomy and to discriminate the high achievers from the low

achievers. If it is so, then such MCQs may be incorporation into future question bank with reliability.

The present study was undertaken to analyse the MCQs with the help of item analysis and select the items which are good enough to assess the knowledge of subject anatomy and to discriminate the high achievers from the low achievers. If it is so, then such MCQs may be incorporation into future question bank with reliability.

If we assess year wise items then in the year 2013 SA1, number of items with acceptable level FV was 40%, SA2 had 40%, 2014 SA1 had 60% and SA2 had 26.6%, in 2015 SA1 had 43.3% while SA2 had 26.6%, in 2016 SA1 had 53% while SA2 had 63% items, SA1 of 2017 had 60% and SA2 had 53% items in acceptable FV range, i.e. between 30-70%. In a study by Chauhan et al. they had mean FV of 57.7 [11]. In a study by Karkal and Kundapur the overall mean FV was 56.64%. In this study 71.09% items found to be good [12]. In a study by Mehta and Mokhasi they found 62% items had acceptable FV which is comparable to 2014 SA1 [13].

In the year 2013, DI was good to excellent (53%) in both SA1 and SA2. The DI was good to excellent in 2014 for SA1 (63%) and SA2 (53%) respectively. In the year 2015 SA1 (67%) and SA2 (40%) items with good to excellent DI. In 2016 there are 50% items with good to excellent DI in both SA1 and SA2 while in 2017 these are 60% and 66% respectively for SA1 and SA2. In a study done by Karkal et al. only 40% Items had good to excellent DI [12]. In a study by Gyata Mehta et al. 35% items were with good to excellent DI [13]. All of our year wise SA1 and SA2 items were having better DI. In a study by Kumar et al 66.66% items had good to excellent DI [14].

According to DE in the year 2013 SA1 functional distractors were 46% and 51%. In the year 2014 they are 69% and 49% respectively. In the year 2015 it is 66% and 43% respectively. In the year 2016 it is 65% and 69% in SA1 and SA2. In 2017 it is 75% and 63% in SA1 and SA2 respectively. In a study by Namdeo et al. the FD were 46.6% which is comparable to 2013 SA1 [15]. According to a study by Patil et al. the functional distractors were 82.2% which is very high [16]. In a study by Tarrant M et al, the proportion of items containing all three functioning distractor was 13.8% [17]. It is better to have an item with two plausible distractors rather than an item with three or four implausible distractors [18]. If we see the correlation coefficient then we find that in the year 2013 it was 0.71 and 0.78 in SA2. In 2014 it is 0.66 and 0.70. In 2015 it is 0.79 and 0.65. In the year 2016 it is 0.64 and 0.64. In 2017 it is 0.42 and 0.17. It indicates good correlation means good reliability of the methods used. In a study by Kumar et al. the correlation co efficient is 0.65 which is comparable to our 2016 SA1, 2 and 2015 SA1 [14].

According to Bloom's classification the formation of MCQs should be for assessing the cognitive, affective and psychomotor domains of the students. We have conducted our study so as to assess whether MCQs are effective to assess the overall knowledge of anatomy but our questions are mainly factual. Our correlation co-efficient in all the five years is positive so we can say that our MCQs are reliable enough to assess the knowledge of anatomy after one year of teaching. Most of the questions are of acceptable range. Non-functional distractors should be reframed and reassessed. Reframing and reassessing is a continuous cycle and should be done yearly till the result becomes consistent and then they should be included in the question bank. The MCQ format allows teachers to efficiently assess deep approach of large numbers of candidates and test a wide range of content [19, 20].

By doing assessment by MCQs not only the knowledge of the students but also the efficiency of teaching process is evaluated. By constant evaluation the poor achievers can be given more time and teaching can be made more effective. We can say that in the year 2015 SA2 had >50% unacceptable discriminative power otherwise all other papers had items with acceptable discriminative in more than 50%. Few year papers like 2013 SA1, 2016 SA1 and 2015 SA1, SA2 and 2014 SA1 had few items with negative discriminative index. So these items should be revised reframed and then reconsidered for future question bank. Overall we can say that the items of the papers have 50% items with good to excellent discriminative index. So we can say that mean distractor effectiveness of all papers except SA2 of 2013

and 14 is functional. We can say that the Kuder-Richerdson correlation coefficient-20 is positive, so the MCQ papers are reliable for the assessment of knowledge of anatomy of the students.

CONCLUSION

This study was based on the data comprising mostly factual knowledge. The framing of MCQ should have also been included other domains such as cognitive and affective. During the validation process, the recognised MCQ having poor and negative DI or having poorly framed stem should be reframed and again subject to revalidation. Those distractors which are non-functional should be replaced by more effective distractors.

ACKNOWLEDGEMENTS

We are thankful to Controller of Examination, Dr. Ramji Singh for providing us the study materials without which it would not be possible to complete the study.

Conflicts of Interests: None

REFERENCES

[1]. Srivastava A, Dhar A, Agarwal CS. Why MCQ? Indian Journal of Surgery2004;66:246-48.
[2]. Carneson J, Delpierre G, Masters K. Designing and managing MCQs. Appendix C: MCQs and Blooms taxonomy. Available at: http://web. uct.ac.za/projects/cbe/mcqman/mcqappc.html. Accessed on 11 December 2017.
[3]. Case SM, Swanson DB. Extended matching items: a practical alternative to free response questions. Tech Learn med 1993;5:107-15.
[4]. Hingorjo MR, Jaleel F. Analysis of one-best MCQs: the difficulty index, discrimination index and distracter efficiency. J Pak Med Assoc. 2012;62:142-6.
[5]. Haladyna TM, Downing SM, Rodriguez MC. A review of multiple-choice item- writing guidelines for classroom assessment. Appl Measurement Edu. 2002:15(3):309-34.
[6]. Gajjar S, Sharma R, Kumar P, Rana M. Item and test analysis to identify quality multiple choice questions (MCQs) from an assessment of medical students of Ahmedabad, Gujarat. Indian J Comm Med. 2014;39(1):17-20.
[7]. Sharif MR, Rahimi SM, Rajabi M, Sayyah M. Computer software application in item analysis of exams in a college of medicine. J Sci Tech. 2014;4(10):565-9.
[8]. Vyas R, Supe A. Multiple choice questions: a literature review on the optimal number of options. The National J India. 2008;21(3):130-3.
[9]. Kehoe J. Writing multiple-choice test items. Pract Assess, Res Evaluation. 2008;4(9). Available online: http://PAREonline.net/getvn.asp?v=4&n=9.
[10]. Singh T, Gupta P, Singh D. Test and item analysis. Principles of Medical Education. 3rd Ed. New Delhi, Jaypee Brothers Medical Publishers (P) Ltd; 2009: 70-77.
[11]. Chauhan P, Chauhan GR, Chauhan RB, Vaza JV and Rathod SP. Relationship between difficulty index and distracter effectiveness in single best-answer stem type multiple choice question. Int J Anat Res. 2015, 3(4): 1607-10.
[12]. Karkal YR and Kundapur GS. Item analysis of multiple choice questions of undergraduate pharmacology examinations in an international medical school in India. 2016. 5(3). 183-186.
[13]. Mehta G and Mokhasi. Item Analysis of Multiple Choice Questions- An Assessment of the Assessment Tool. 2014. 4(7). 197-202.
[14]. Kumar T, Kumar Y, Jha K and Singh R. Adding Value to Multiple Choice Questions Banking: An AIIMS Patna Experience. Indian J Physiol Pharmacol. 2017. 61(1): 70-75.
[15]. Namdeo SK, Sahoo B. Item analysis of multiple choice questions from an assessment of medical students in Bhubaneswar, India. Int J Res Med Sci. 2016. 4(5):1716-1719.
[16]. Patil R, Bhaskar SP, Vell K, Boratne AV. Evaluation of multiple choice questions by item analysis in a medical college at Pondicherry, India. Int J Community Med Public Health. 2016; 3(6):1612-16.
[17]. Tarrant M, Ware J, Mohammed AM. An assessment of functioning and non-functioning distractors in multiple-choice questions: A descriptive analysis. BMC Medical Education. 2009;9:1-8.
[18]. Schuwirth LWT, Vleuten CPM. Different written assessment methods: what can be said about their strengths and weaknesses? Medical Education. 2004;38(9):974.
[19]. Downing SM: Assessment of knowledge with written test forms. In International handbook of research in medical education Volume II. Edited by:Norman GR, Van der Vleuten C, Newble DI. Dorcrecht: Kluwer Academic Publishers. 2002; 647-72.
[20]. McCoubrie P. Improving the fairness of multiple choice questions: a literature review. Med Teach. 2004; 26(8):709-12.

How to cite this article:

Padamjeet Panchal, Bheem Prasad, Sarita Kumari. MULTIPLE CHOICE QUESTIONS - ROLE IN ASSESSMENT OF COMPETENCY OF KNOWLEDGE IN ANATOMY. Int J Anat Res 2018;6(2.1):5156-5162. **DOI:** 10.16965/ijar.2018.143