Analysis of Difficulty Index, Discrimination Index and Distractor Efficiency of Multiple Choice Questions of Speech Sounds of English

Lok Raj Sharma

Associate Professor of English

Makawanpur Multiple Campus, Hetauda, Nepal

E-mail:lokraj043@gmail.com

ABSTRACT

The prime purpose of this research article is to analyze 20 multiple choice questions (MCQs) regarding the consonant and vowel sounds of English having asked to the Bachelor of Education (B. Ed) first year major English students in their achievement examination conducted at Makawanpur Multiple Campus, Hetauda, Nepal in 2020. The researcher employed 33% students from the group of high achievers from the top and 33% students from the group of low achievers from the bottom by including 18 students from the population of 27 students. Each item was analyzed for difficulty index (DIF I), discrimination index (DI), and distractor effectiveness (DE). 3 (15%) questions fell in the range of the difficulty index of (0.20 – 0.39). 14 (70%) questions fell in the range of the difficulty index of (0.40 – 0.59). 2 (10%) questions fell in the range of the difficulty index of (0.60 – 0.79). 1 (5%) Question fell in the range of the difficulty index of (0.80 – 0.89). 2 (10%) questions lay in the discrimination index of (0.20-0.29) and 18 (90%) questions lay in the discrimination index of ≥ 0.40. 5 (25%) questions had one non-functional distractor, whereas 15 (75%) questions had zero nun-functional distractor. Out of 60 distractors, 5 (8.34%) distractors were non-functional and 55 (91.66%) distractors were functional. It was found that most of the multiple choice questions were reliable and valid.

Keywords: Difficulty index, discrimination index, distractor effectiveness, item analysis, multiple choice questions.

INTRODUCTION

Item analysis is a process of collecting, summarizing, and using information from students' responses to assess the quality of multiple-choice questions (MCQs). Performing an item analysis is a prominent aspect of maintaining the quality of multiple choice questions asked to the students. It is based on the responses given by the examinees. The most difficult questions or the easiest questions need to be discarded or revised. The decision to revise the questions is based on good difficulty index, discriminatory index and functional distractors.

Multiple choice questions are very common objective tests used in the faculty of education in Nepal. Such questions require the students to select one answer from the list of alternatives in which only one alternative is correct, whereas other alternatives are wrong. The correct alternative is called the key and the wrong alternatives are called distractors. The multiple choice questions are implemented in entrance examinations of medical sector, engineering, science, civil services and so on. They are compulsorily asked to the students in the Faculty of Education in the internal and final examinations in Nepal.

Teaching speech sounds to the Nepalese students is a difficult task, but they are compelled to learn because their course of study includes speech sounds in written and practical exams. The article writer first taught speech sounds of English to the students and administered achievement test through 20 multiple choice questions based on the speech sounds of English for the purpose of writing a journal article. Each question involves 4 alternatives, where 1 alternative was the correct/ right answer and 3 alternatives were the wrong answers or distractors. It means there were 60 distractors in all. The writer was curious to examine whether the questions were valid and reliable in principle or not. The general standard involves taking 27% students from the group of high achievers from the top and 27% students from the group of low achievers from the bottom. But he took 33% from both groups, because the number of students was low. The sample consists of 18 students from the population of 27 students.

I was interested in this field, because it is the field in which most of the teachers and the question setters pay little attention in spite their continual involvement in the work. We can find a few researches regarding the analysis of multiple choice questions in medical sectors in India. The writer considers that it is a novel small step towards writing on item analysis in Nepal. This article is helpful to those who are involved in the field of teaching learning as well as question setting activities. It is dedicated to those who are concerned with the quality of the MCQS prepared for the examinees to assess their understanding of particular subjects.

OBJECTIVE OF THE STUDY

The key objective of the research study is to:

♦ Analyze the difficulty index and discrimination index and distractor efficiency in multiple-choice questions based on speech sounds of English.

DELIMITATIONS OF THE STUDY

The research study was carried out under the following delimitations:

♦ The study included only 18 students as the respondents.
♦ Only B.Ed. first year students majoring English were involved in the study.
♦ The study was delimited to only one campus.
♦ Each multiple choice question was constructed with one key and three distractors.

THEORETICAL LITERATURE REVIEW

Item Analysis

Item is a statement in the form of a question. Item analysis is one of the most important aspects of test construction. Item analysis is a general term for a set of methods used to evaluate test items. Items can be analyzed qualitatively in terms of their content and form and quantitatively in terms of their statistical properties. One of the ways of evaluating the students' understanding of subject matters is through an assessment that is an essential phenomenon in the course of teaching and learning activities. Popham (2002) and Trice (2000) state that students' assessment and evaluation are an integral part of the teaching and learning process Assessing the students' understanding can be done through test items.

Item analysis is a statistical technique which is used for selecting and rejecting the items of the test on the basis of their difficulty index and discrimination index. Item Analysis is an important tool to increase the effectiveness of the test. It is a process which examines student's responses to individual test items to assess the quality of these items and quality of test as a whole.

Gronlund (1993) asserts that item analysis allows us to observe the item characteristics, and to improve the quality of the test. Lange and Mehrens (1967) consider that item revision allows us to identify items too difficult or too easy, items not able to differentiate between students who have learned the content and those who have not, or questions that have distractors not plausible. Singh, Gupta and Singh (2009) hold the view that item analysis as a process which assesses the quality of those items and of the test as a whole. Zubairi and Kassim (2006) and Sim and Rasiah (2006) affirm that item analysis provides feedback to teachers for necessary modifications in MCQs to make it suitable for the exam. While some MCQs are edited, some are deleted based on the analysis. Sharma (2000) and Freeman (1962) consider that the quality of a test depends upon each items of a test. Considine, Botti and Thomas (2005) believe that 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.

Multiple Choice Questions

A MCQ is composed of a stem and several options. The question or the statement in the sense of a question is called the stem. The correct option is called the key while the incorrect alternatives are called the distractors. MCQs are used mostly for comprehensive assessment at the end of academic sessions and they provide feedback to the teachers on their educational actions. Cizek and O'Day (1994) affirm that MCQ consists of a stem with a question followed by number of options. One of the options is the best or correct response known as the key while others are described as distractors.

Brown (2004) considers that multiple-choice items are described as receptive or selective. Öztürk (2007) highlights the importance of multiple choice items and remarks   that multiple-choice items seem to be reliable compared with other types of tests which are negatively affected by subjectivity. Multiple choice-items were significantly easier and less discriminating than free

response tasks. Hughes (2003) emphasizes some weaknesses of multiple-choice items stating that this technique only tests recognition knowledge which is a lower mental skill according to Bloom's taxonomy (1956).

Difficulty Index (DIF I)

The difficulty index (DIF I) of an item is defined as proportion or percentage of the examinees who correctly answered a given test item. It is one of the key parameters of item analysis. It ranges from 0 to 1 or (0% to 100%). We can use the following formula to calculate the difficulty Index.

$$\text{DIF I} = \frac{RH+RL}{NH+NL} \quad \text{or} \quad = \frac{RH+RL}{N}$$

Where, RH→The number of right answers in the higher group; RL→The number of right answers in the lower group; NH→The number of examinees in the higher group; N→Total number of examinees and NL→ The number of examinees in the lower group.

Discrimination Index (DI)

Index of discrimination is that ability of an item on the basis of which the discrimination is made between the superior (the group of high achievers) and the inferior (the group of low achievers). It ranges from -1 to +1. Gajjar, Sharma, Kumar and Rana (2014) define item discrimination as "the ability of an item to differentiate between students of higher and lower abilities" (p.18).

We can use the following Formula for determining the discrimination Index:

$$\text{DI} = \frac{2\ (RH-RL)}{N}$$

Distractor

A distractor is a wrong option in the multiple choice question. It is designed to see whether the person being tested can notice the difference in a test or not. A distractor is considered to be a good distractor when it attracts more examinees from the group of low achievers than the group of high achievers. According to Malau-Aduli and Zimitat (2012), a distractor that fails to "attract any examinees is dysfunctional, does not assist in the measuring of educational outcomes, adds nothing to the item or the test (psychometrically) and has negative impact upon learners" (p.927). Mehta and Mokhasi (2014) assert that distractor are important components of an item and have a great impact on the total test score. Student's performance depends on how the distractors are designed.

There are two types of distractor. They are: non-functional distractor (NFD) and functional distractor (FD). Non-functional distractor (NFD) in an item is the option, other than the correct option selected by less than 5% of students and the functional or effective distractor is the option selected by 5% or more. Tarrant, Ware and Mohammed (2009), Vyas, and Supe (2008), and Patil and Patil (2015) hold their view that functional distractors (FD) are those that are selected by >5% or more of the examinees and non-functional distractors (NFD) are the options selected by <5% of the examinees.

The percent of a distractor can be calculated by using the following formula:

$$\textbf{Percent of a distractor} = \frac{\text{Number of examinees of the Distractor}}{\text{Total Number of examinees}} \times 100\ \%$$

Effectiveness of Distractors / Distractor Efficiency (DE)

Distractor efficiency (DE) for any item ranges from 0 to 100% and is determined on the basis of the number of NFDs in an item. It means DE is expressed as 0%, 33.3%, 66.6% and 100% depending on the number of NFDs.

Number of Non- functional Distractors	Distractor Efficiency (DE)
3 NFDs	0.00 %
2NFDs	33.33 %
1NFD	66.66 %
0 NFD	100 %

Empirical Literature Review

A research study carried out by Burud, Nagandla and Agarwal (2019) using 120 multiple choice questions administered to 113 medical students at International Medical University, Malaysia, shows that 35 (29.16%) were very easy questions, 18 (15%) , 42 (35%)were excellent, 20 (16.67 %) were good and 5 (4.17 %) were very difficult questions. With respect to discriminatory index, 25 (20.83%) items showed excellent difficulty index, 35 (29.17%) items showed good discrimination, 37 (30.83%) items showed fair discrimination, 13 (10.83%) items showed poor discrimination and 10 (8. 34%) items showed negative discrimination. Regarding the distractor efficiency, 47 (39.16%) items had no NFD, 51 (42.5%) items had 1 NFD 18 (15.0%) items had 2 NFDs and 4 (3.34%) items had 3 NFDs.

A research study accomplished by Mahjabeen, Alam, Hassan, Zafar, Butt, Konain and Rizvi (2018) carried out a research study with 65 multiple choice questions administered to 110 4th-year MBBS students at Islamabad Medical and Dental College Islamabad during the academic session 2017 shows that 53(81%)
MCQs were in acceptable category, only 1(2%) MCQ was too difficult and 11(17%) were too easy. Regarding DI, total 34(62%) MCQs showed excellent discrimination tendency to distinguish low and high performer students. While 15(23%), 5(8%) and 11(17%) MCQs demonstrated good, acceptable and poor discrimination ability respectively. Out of total 260 distractors, 72% were functional and only 28% were non-functional. Total 16(25%) MCQs had zero non-functional distractor (NFDs), while 30(46%) and 16(25%) MCQs had 1 and 2 NFDs respectively. Only 3(5%) MCQs were with 3 or more NFDs.

Namdeo and Rout in their study (2016) having 25 multiple choice questions administered to 100   MBBS students at Kalinga Institute of Medical Science, Bhubaneswar shows that total 25 MCQs and 75 distractors were analyzed. Out of 75 distractors, 40 (53.4%) NFDs were present in 22 items. 3(12%) items had no NFDs while 8(32%), 10(40%), and 4 (16%) items contained 1, 2, and 3 NFDs respectively.

Patil, Palve1, Vell and Boratne (2016) carried out a study with 30 multiple choice questions administered to 22 medical students at Mahatma Gandhi Medical College and Research Institute, Pillaiyarkuppam, Sri Balaji Vidyapeeth University, Pondicherry. Total 90 distractors (3x30 MCQs) were analysed. Mean for difficulty index, discrimination index and distractor efficiency were 38.3%, 0.27 and 82.8% respectively. Of 30 items, 11 items were of higher difficulty level (DIF I <30%) while 5 items were of easy level (DIF I >60%). Total 15 items were having very good DI. Of the 90 distractors, there were 16 (17.8%) non-functional distractors (NFDs) present in 13 (43.3%) items.

A research study done by Namdeo and Sahoo (2016) using 25 multiple choice questions administered to 76 medical students at Kalinga Institute of Medical Science (KIMS) Bhubaneswar shows that   Difficulty index of 14 (56%) items was in the acceptable range (p value 30-70%), 8 (32%) items were too easy (p value >70%) and 2 (8%) items were too difficult (p value <30%). Discrimination index of 12 (48%) items was excellent (d value>0.35), 3 (12%) items was good (d value 0.20-0.34) and 8(32%) items were poor (d value<0.2%). Out of 75 distractors, 40 (53.4%) NFDs were present in 22 items. 3 (12%) items had no NFDs, whereas 8 (32%), 10 (40%), and 4 (16%) items contained 1, 2, and 3 NFD respectively.

A research study carried out by Angadi, Nagabhushana and Hashilkar (2018) using 50 multiple choice questions administered to 180 J. N. Medical College Students, Belagavi, Karnataka, A total of 50 items consisting of 150 Distractor s were analysed. DIF I of 31 (62%) items were in the acceptable range (DIF I= 30-70%) and 30 had 'good to excellent' (DI >0.25). 10 (20%) items were too easy and 9 (18%) items were too difficult (DIF I <30%). There were 4 items with 6 non-functional Distractor s (NFDs), while the rest 46 items did not have any NFDs.

A research study done by Rao, Kishan, Sajitha, Permi, and Shetty (2019) with 40 multiple choice questions administered to 120 2nd year MBBS students at K. S. Hegde Medical Academy, , Karnataka, India . The study conducted in the Department of Pathology shows that difficulty index of 34 (85%) items was in the acceptable range (P = 30–70%), 2 (5%) item was too easy (P >70%), and 4 (10%) items were too difficult (P <30%). Discrimination index of 24 (60%) items was excellent (D >0.4), 4 (10%) items were good (D =0.3–0.39), 6 (15%) items were acceptable (D =0.2–0.29), and 6 (15%) items were poor (D < 0–0.19). A total 40 items had 120 distractors. Amongst these, 6 (5%) were nonfunctional distractors, 114 (95%) were functional distractors. The discrimination index exhibited positive correlation with difficulty index (r = 0.563, P = 0.010, significant at 0.01 level [two-tailed]). The maximum discrimination (D = 0.5–0.6) was observed in acceptable range (P = 30–70%).

RESEARCH METHODOLOGY

Research Design

A cross-sectional survey design was used to carry out the study. The researcher collected the primary data from the questionnaire to analyze the difficulty index, discrimination index and distractor efficiency in multiple choice questions asked to the Bachelor of Education first year students in their internal achievement examination.

Population / Universe

The population of the study consisted of 27 Bachelor of Education students studying major English at Makawanpur Multiple Campus, Nepal in the Academic Year 2020.

Sampling Design

The researcher selected 18 students by following the standard of selecting 33% respondents from the group of high achievers from the top and 33% respondents from the group of low achievers from the bottom after maintaining the scores in descending order.

Sample Size by Gender

The sample size of the research study consisted of 5 boys and 13 girls studying at the campus.
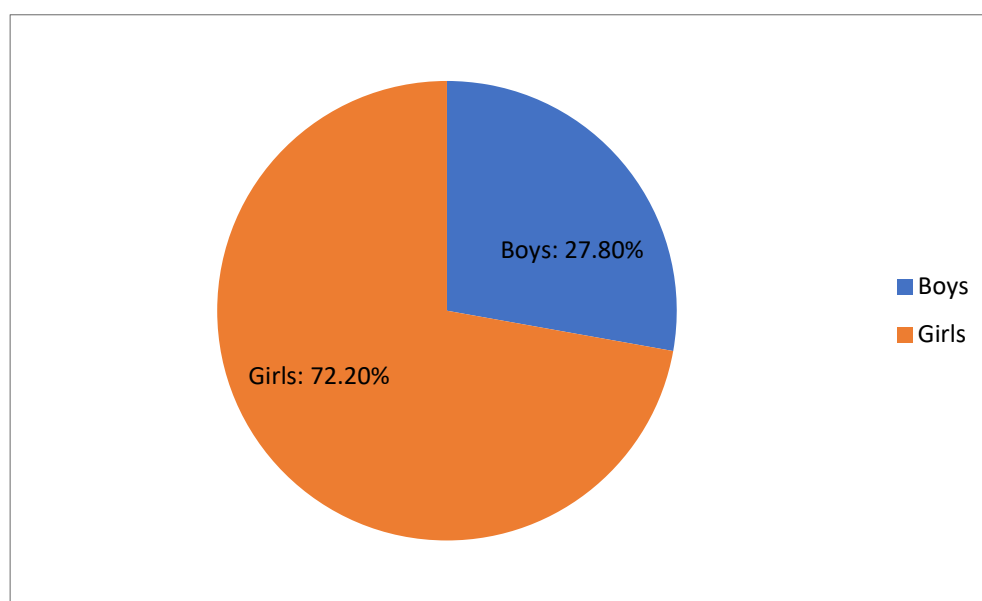

Figure 1: Percent of students by gender

This figure shows that the percent of boys was a smaller than that of the girls.

Variables in the Study

The researcher took twenty multiple choice items with four alternatives (1 correct answer and 3 distractors in each and every question) as major variables. Other variables that are analyzed are difficulty index, discrimination index and distractor efficiency.

Validity and Reliability of the Instruments

The validity of questionnaire was conducted by using Pearson Product moment correlations in SPSS, where the score of each item was correlated with the total score. The Pearson Product moment correlation count value was greater than the corresponding Pearson correlation critical value $|r| > rc = 0.422$ of all questions except question numbers 18 and 20. It indicates that most of the questions were valid. The Split-half reliability test was used to check the reliability of the questions. The overall internal consistency of questions based on Guttman Split-Half Coefficient was 0.969 which indicates that the questions were highly reliable.

Nature and Source of Data

The researcher exploited ratio data to carry out the research study. Students' performance in speech sounds of English was tested through the use of objective questions. The primary source of data was questionnaire. The secondary source of data included books, journal articles, web-sites etc.

Analysis and Interpretation of Data

All the data were analyzed by using Statistical Package for Social Sciences (SPSS) 20 version. The researcher used a series of twenty multiple choice questions asked in the internal achievement examination to the Bachelor of Education first year students to analyze the test items. The researcher used the frequency and percent statistics to analyze the test items by focusing the difficulty index, the discrimination index and the effectiveness of distractors in each test item.

Table 1: Number of examinees who answered rightly in both groups

Item No.	Right answers from the group of high achievers (RH)	Right answers from the group of low achievers (RL)	Total right answers	Item No.	Right answers from the group of high achievers (RH)	Right answers from the group of low achievers (RL)	Total right answers
1	8	1	9	11	9	1	10
2	4	0	4	12	7	0	7
3	7	1	8	13	7	1	9
4	7	2	9	14	8	2	10
5	7	2	9	15	7	2	9
6	8	1	9	16	8	2	10
7	5	0	5	17	6	2	8
8	7	1	8	18	8	6	14
9	7	2	9	19	8	3	11
10	7	2	9	20	9	7	16

Calculation of Difficulty Index and Discrimination Index

The difficulty index and the discrimination index of each question item can be computed by using the following formulae:

Difficulty Index (DIF I) $= \frac{RH+RL}{N}$ &   Discrimination Index (DI) $= \frac{2\,(RH-RL)}{N}$

Table 2: Difficulty index, discrimination index and number of non-functional distractors

MCQ Item Number	Difficulty Index (DIF I)	Discrimination Index (DI)	Number of NFDs	Effectiveness of Distractors (EDs) in (%)
1	0.50	0.77	0	100.00
2	0.22	0.44	0	100.00
3	0.44	0.66	1	66.66
4	0.50	0.55	0	100.00
5	0.50	0.55	0	100.00
6	0.50	0.77	0	100.00
7	0.27	0.55	0	100.00
8	0.44	0.66	1	66.66
9	0.50	0.55	0	100.00
10	0.50	0.55	0	100.00
11	0.55	0.88	0	100.00
12	0.38	0.77	0	100.00
13	0.44	0.66	1	66.66
14	0.55	0.66	0	100.00
15	0.50	0.55	0	100.00
16	0.55	0.66	0	100.00
17	0.44	0.44	0	100.00
18	0.77	0.22	1	66.66
19	0.61	0.55	0	100.00
20	0.88	0.22	1	66.66

4.2 Interpretation of Difficulty Index

The researcher has employed the following table to analyze the difficulty index of each question item.

Table 3: Evaluation of difficulty index

S.N.	Difficulty Index	Number of Items	Item Evaluation	Recommendation
1.	< 0.20	0	Most difficult	Remove
2.	0.20-0.39	3	Difficult	Keep
3.	0.40-0.59	14	Moderately difficult	Keep
4.	0.60-0.79	2	Moderately Easy	Keep
5	.80-. 89	1	Easy	Keep
6.	> 0.90	0	Easiest	Remove

Question items 2, 7 and 12 fell in the range of the difficulty index of (0.20 – 0.39). It shows that 3 questions were difficult. Question items 1,3, 4, 5, 6,8, 9, 10, 11, 13, 14, 15, 16 and 17 fell in the range of the difficulty index of (0.40 – 0.59). It indicates that 14 questions were moderately difficult. Question items 18 and 19 fell in the range of the difficulty index of (0.60 – 0.79). Those 2 questions were moderately easy. Question item 20 fell in the range of the difficulty index of (0.80 – 0.89). This question was easy. The table shows the questions should be maintained in the list of questionnaire.

Interpretation of discrimination index

The researcher used the following table to analyze and interpret the discrimination index of each question item.

Table 4: Evaluation of discrimination index

S.N.	Discrimination Index	Number of Items	Percent	Evaluation of Item	Recommendation
1.	Negative	0	0	Worst/ defective item	Definitely Discard
2.	< 0.20	0	0	Not discriminating item, marginal item	Revise / Discard
3.	0.20- 0.29	2	10	Moderately discriminating, fair item	Keep
4.	.30- 0.39	0	0	Discriminating item, good item	Keep
5.	≥ 0.40	18	90	Very discriminating, very good item	keep

This table demonstrates question items 18 and 20 that lay in the discrimination index of (0.20-0.29) were moderately discriminating items and were acceptable for the inclusion in the list. All the question items except the two questions were greater than 0.40 of the discrimination index. They were very discriminating items, excellent items that must have been kept in the list.

Calculation of the Effectiveness of Distractors / Distractor Efficiency in Each Item

Distractor efficiency is determined on the basis of the number of NFDs in an item. It means DE is expressed as 0%, 33.3%, 66.6% and 100% depending on the number of NFDs. The percent of each distractor was calculated by using the following formula:

$$\textbf{Percent of a distractor} = \frac{\textbf{Number of examinees of the Distractor}}{\textbf{Total Number of examinees}} \times 100 \,\%$$

Table 5: Distractor efficiency in each item

Number of Non-functional Distractors	Question No.	Distractor Efficiency (DE)
1NFD	3,8,13,18,20	66.66 %
0 NFD	1,2,4,5,6,7,9,10,11,12,14,15,16,17,19	100 %

There were 60 distractors in the question items. 55 (91.67%) distractors were functional and 5 (8. 33%) distractors were non-functional. It means most of the distractors were effective in test items. In other words, 5 questions had one non-functional distractors, whereas 15 had functional distractors. It can be said that the effectiveness of distractors in 5 questions was 66.66% and that of in 15 questions was 100.00%.

Calculation of Mean Scores of the Difficulty Index and the Discrimination Index

The researcher computed the mean scores of the difficulty index and the discrimination index to evaluate the items as a whole.

Table 6: Mean and standard deviation of difficulty index

Aspect	No. of Items	Range	Minimum	maximum	Mean	Standard Deviation
Difficulty Index	20	.66	.22	.88	.50	.14

The mean score of the difficulty index was 0.50 that lay between 0.40- 0.59. It shows that the test items were moderately difficult and they should be kept in the list of test items.

Table 7: Mean and standard deviation of discrimination index

Aspect	No. of Items	Range	Minimum	maximum	Mean	Standard Deviation
Discrimination Index	20	.66	.22	.88	.58	.16

This table shows that the mean score of the discrimination index was 0.58. It reveals that all the multiple choice questions were excellent from the perspective of the standard of discrimination index value. It suggests that the examiner should keep these items in the list of the test.

RESULTS

All the multiple choice questions fell in the difficulty index range between 0.20-0.89 and that was the range of accepting the test items. 3 questions fell in the range of the difficulty index of (0.20 – 0.39). It shows that those questions were difficult. 14 questions fell in the range of the difficulty index of (0.40 – 0.69). It indicates that those questions were moderately difficult. 2 questions fell in the range of the difficulty index of (0.60 – 0.79). Those questions were moderately easy. 1 question fell in the range of the difficulty index of (0.80 – 0.89). That question was easy. The mean score of the difficulty indices of the test items was 0.50. It indicates that the questions were moderately difficult in all. Similarly, 2 question items that lay in the

discrimination index of (0.20-0.29) were moderately discriminating items and were acceptable for the inclusion in the list. 18 question items that lay in ≥ 0.40 the discrimination index were very discriminating items, excellent items that must have been maintained in the list. The mean score of the discrimination indices of the question items was 0.58 and the scale shows that question items were very discriminating and excellent. 55 (91.67%) distractors were functional, whereas 5 (8.33%) distractors were non-functional.

CONCLUSION

Item analysis is a prominent procedure performed after the examination for providing information regarding the reliability and validity of an item/test by calculating difficulty index, discrimination index and distractor efficiency. In this study, the majority of question items fulfilled the criteria of acceptable difficulty and good discrimination which means the MCQs selected were of good quality. The researcher would like to recommend the other researchers to carry out further item analyses after each examination to identify the areas of potential weakness in the formation of MCQ items to improve the standard of assessment of examinees' understanding of subject matters.

ACKNOWLEDGEMENTS

REFERENCES

1. Angadi, N. B., Nagabhushana, A., & Hashilkar, N.K. (2018). Item analysis of multiple choice questions of undergraduate pharmacology examinations in a medical college in Belagavi, Karnataka, India. *International Journal of Basic & Clinical Pharmacology, 7*(10). 1917-1920.

2. Brown, H. D. (2004). *Language assessment: Principles and classroom practices.* White Plains, NY: Pearson Edu¬cation.

3. Burud, I., Nagandla, K., & Agarwal, P. (2019). Impact of distractors in item analysis of multiple choice questions. *International Journal of Research in Medical Sciences, 7* (4), 1136-1139.

4. Cizek, G.J., & O'Day, D.M. (1994). Further investigation of nonfunctioning options in multiple-choice test items. *Educational Psychol Measurement*, *54*(4),861-872.

5. Considine, J., Botti, M., & Thomas, S. (2005). Design, format, validity and reliability of multiple choice questions for use in nursing research and education. *Collegian. 12* (1),19-24.

6. Freeman, F. (1962). *Theory and practice of psychological testing.* New Delhi: Oxford & Ibh publishing.

7. Gajjar, S., Sharma, R., Kumar, P., & Rana, M. (2014). Item and test analysis to identify quality multiple choice questions (MCQs) from an assessment of medical stu¬dents of Ahmedabad, Gujarat . *Indian Journal of Com¬munity Medicine, 39*(1), 17-22.

8. Gronlund, N. E. (1998). *Assessment of student achievement. 6th edition.* Boston: Allyn and Bacon.

9. Hughes, A. (2003). *Testing for language teachers* (2 nd ed.). Cambridge: Cambridge University Press.

10. Lange, A., Lehmann, I.J. & Mehrens, W.A. (1967). Using item analysis to improve tests. *Journal of Educational Measurement, 4*(2), 65-68. Retrieved from http://www.jstor.org/stable/1434299.

11. Mahjabeen, W., Alam, S., Hassan, U., Zafar, T., Butt, R., Konain, S., Rizvi, M. (2017). Difficulty index, discrimination index and distractor efficiency in multiple choice questions. *Annals of PIMS,* 310-315.

12. Malau-Aduli, B. S., & Zimitat, C. (2012). Peer review im¬proves the quality of MCQ examinations . *Assessment & Evaluation in Higher Education, 37*(8), 919-931.

13. Mehta, G., & Mokhasi, V. (2014). Item analysis of multiple choice questions: An assessment of the assessment tool. *International Journal Health Science Research, 4*,197-202.

14. Namdeo, S.K., & Rout, S.D. (2016). Assessment of functional and nonfunctional distractor in an item analysis. *International Journal of Contemporary Medical Research,* 3 (7),1891-1893.

15. Namdeo S.K.,& Sahoo B. (2016).Item analysis of multiple choice questions from an assessment of medical students in Bhubaneswar, India. *International Journal of Research in Medical Sciences*, *4* (5),1716-1719.

16. Öztürk, M. (2007). Multiple-choice test items of foreign language vocabulary. *Eğitim Fakültesi Dergisi, 20*(2), 399-426.

17. Patil, R., Palve, S.B., Vell, K., & Boratne, A.V. (2016). Evaluation of multiple choice questions by item analysis in a medical college at Pondicherry, India. *International Journal of Community Medicine and Public Health*, *3(*6),1612-1616.

18. Patil V. C, & Patil, H. V. (2015). Item analysis of medicine multiple choice questions (MCQs) for under graduate (3rd year MBBS) Students. *Research Journal of Pharmaceut Biol Chem Sci*, *6*,1242-1251.

19. Popham, W. J. (2008). *Classroom assessment: What teachers need to Know (5th ed.).* Boston: Allyn and Bacon.

20. Rao, C., Kishan Prasad, HL., Sajitha, K., Permi, H., & Shetty J. (2019). Item analysis of multiple choice questions: Assessing an assessment tool in medical students. *International Journal of Educational and Psychological Researches, 2* (4)*,* 201-204.

21. Sharma, S. (2008). *Modern teaching strategies. New Delhi: Omsons Publications.* New Delhi: Omsons Publications.

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

23. Singh T, Gupta P, Singh D. (2009). Test and item analysis . *Principles of Medical Education 3rd ed*, 70-77.

24. Tarrant, M., Ware, J., & Mohammed, A.M. (2009). An assessment of functioning and non-functioning distractors in multiple choice questions: a descriptive analysis. *BMC Medicine Education*, *9*(40),2-8.

25. Trice, A. D. (2000). *A handbook of classroom assessment.* New York: Longman.

26. Vyas, R., & Supe, A. (2008). Multiple choice questions: A literature review on the optimal number of options. *National Medical Journal, India*, *21,* 130-133.

27. Zubairi A.M. & Kassim N.L. (2006). Classical and Rasch analysis of dichotomously scored reading comprehension test items. *Malaysain Journal of English Language Teaching Research, 2*, 1-20.