Archives of Educational Studies
Vol. 2, No. 2, Dec., 2022, 205-222
https://ares.pk/

Item Analysis of Tool used for Examining the Effectiveness of E-modules for Academic Performance of 7th Grade Students

Arshad Mehmood Qamar[1*], Dr. Wajiha Kanwal[2] & Rashda Perveen[3]

[1]Lecturer Science Education Department, Allama Iqbal Open University Islamabad, Pakistan. E-Mail: arshad.mehmood@aiou.edu.pk
[2]Assistant Professor Department of Education, University of Wah, Rawalpindi, Pakistan. E-Mail: dr.wajiha@uow.edu.pk
[3]Secondary School Teacher Punjab. E-Mail: ra76.islamabad@gmail.com

Abstract

The main objective of this research study was to conduct analysis of test items for measuring quantitative characteristics (difficulty level, distractor efficiency & discrimination index) of research tool developed from the four units of general science of seventh (7th) class to be used as a test/tool for the research on "examining the effectiveness of e-modules for academic performance of students studying in 7th grade". Tool was developed from the prescribed course of 7th class general science (Chemistry section). Data was collected from 240 students of 7th class. Researchers followed the procedure used by (Qamar, Kanwal & Nadeem, 2022). Items were analyzed for finding difficult level, discrimination index and distractor efficiency. Based on findings two test items (17 & 18) were eliminated & five (were revised because they have very small value of Diff. I (<29) and five test items (2, 17, 21, 22, 23) were revised due to less discrimination index. Thirteen distractors were revised and improved. As a result of item analysis five test items were eliminated & eleven items were refined and retained for data collection. Overall twenty five items could be used for data collection.

***Keywords**: Item Analysis, Multiple Choice Items (MCQs), Difficulty Index (Diff. I), Discrimination Index (D I), Distractor effectiveness (DE)*

Introduction

Item analysis is a technique used to assess and evaluate the performance and quality of the developed test items in the form of MCQs. Learning cannot be enhanced without integrating assessment with teaching and learning process. The Assessment techniques used in Pakistan include Multiple choice test items, short questions, structured examinations based on SLOs. The potential to use any type of assessment is embedded in structure of curriculum and required competencies. Multiple choice test items is a type of assessment based on the hard work of teachers which make the students easy in attempting the answers. As suggested by Elgadal & Mariod (2021) MCQ items has the capability to test the students' different abilities such as problem solving, interpretation of data, cognitive thinking, critical thinking and curricular competencies in a very short time. But there are some constrains about item analysis. The results of research conducted by Elgadal & Mariod (2021) depicted that the item analysis in multiple choice test items has capability to measure the validity, reliability, its discriminatory efficiency, and technical anomalies for test item development. It was concluded that item analysis was able to determine the characteristic that why students under-performed and help in determining the root causes of this under performance. After improving the root causes effective and precise students' evaluation regarding competencies is ensured.

There is a range of abilities which are evaluated in psychometric domain but here in item analysis only three domains (items difficulty level, Discrimination index and distractor effectiveness) are evaluated. (Atalmiş & Kingston, 2017). Further a question arises why we do item analysis. In answer to this question, we can measure the quality of teaching, understanding of teachers through students' answers regarding test items and parameters of options and answers given at the end of each question statement. Selection of options and answers by the students provide the basis of item analysis. In the view of Qamar, Kanwal & Nadeem (2022) as a result of conducting item analysis the use of this activity was to bring improvement in test items by revising, improving or discarding the items or distractors. Each one follows some underlying rules or principals. Most of the Pakistani schools, colleges and boards follow normal curve or

standard criteria for difficulty level, discrimination index and distractor efficiency for any test. In the same manner in experimental research achievement tests are used for measurement of academic achievement before and after the intervention. Item analysis helps in ensuring the criterion of the valid test items. According to standard normal curve 68.13% lies in the center of the curve whereas 15.87% lie on the left side and 15.87% lie at the right side of the normal curve. If results of the item analysis deviate from the said criteria, it becomes compulsory for the researcher or teacher to revise the test items. Most difficult items may be revised as a result of inferences made on the part of teacher, curriculum and student of structure of the test item.

This analysis was carried out for pretest and posttest for determining the effect of e-modules on the academic achievement of 7th class students studying general science. Teaching of general science was difficult to teach, and students reported that some concepts were very difficult to assimilate. Students had to study general science because general science was compulsory at elementary level. General science at middle level is a type of integrated curriculum with integration of Biology, Chemistry, Physics and Astronomy. Thirty (30) test items were developed from the section of Chemistry. Two test items were removed as a result of expert opinion. Twenty-eight (28) test items were used for item analysis. Item analysis is interesting for researchers because most of the science teachers and test item developers are less conscious regarding this process. There is one another reason that many developers rely on their expertise and experience. The work regarding item analysis is considerable in the fields of medical, engineering and higher education. But in Pakistan at elementary level this work (item analyses) is lacking. Researchers considered that it will be a novel work towards validity of test for excellent results to be used for research and future researchers.

Objectives

Objectives of this research study were to:

1. find out the difficulty level (Diff.I) of the MCQs test items of general science (Chemistry Section).
2. find out the discrimination index (D.I) of the MCQs of General Science (Chemistry Section).

3. find out the distractor efficiency (D.E) of the test items of general science (Chemistry Section)

Review of Literature

Item Analysis

The researchers tried to improve data collection instruments through a number of statistical techniques and tools. Test is a tool which is used to assess the achievement and performance of students for a particular content and subject. A test is collection of different test items. Item is a form of question where answer of a statement is sought through helping students by giving them clue in the form of options (Sharma, 2021).Out of these options one option is correct and other three options are distractors. Item analysis help the item developers to make the test items more useful and valuable. Item analysis is a system and set of systematic procedures used to evaluate the test items for their effective use in data collection. Items analysis can be analyzed in qualitative form where form and structure of the item statement, answer and distractors are examined through experts and experienced persons in relevant fields. Quantitative analysis was done by using statistical procedures to calculate their statistical properties.

As Popham (2002) &Trice (2000) stated that measurement and evaluation of students in certain area is an integral part of the learning and teaching. This assessment can be done through MCQs where items have power to measure the abilities of students for which MCQs have been developed. Item analysis is one of the best method to make the items valid and reliable. Item analysis can be done in three ways viz; difficulty index, discrimination index and distractor analysis. Items can be rejected, accepted or improved through difficulty index and discrimination index whereas distractors are selected or rejected on the basis of distractor analysis (Sharma, 2021).Distractor analysis is a process where we examine the students‟ responses in an individual test item. If an option is not selected by any student, it is assumed that option has no relevancy and hence is nonfunctional. As a result to include maximum distractors (options) which are functional?

According to Gronlund (1993), item analyses make us able to determine the item characteristics and for improving the quality of items.

Sim and Rasiah (2006), Zubairi and Kassim (2006) reported that through results of item analysis, teachers can make necessary changes in test items for making items effective for measurement of achievements of students during examinations Gupta , Singhand Singh (2009) viewed that item analysis is a process to assess the quality of test items as a whole. Botti, Considineand Thomas (2005) believed that test item analyses provide necessary information regarding validity and reliability of a test item.

Multiple Choice Items

Multiple choice items assessment comprises a stem (statement) and many options. Most probably each statement consists of three to five options depending upon the model used and objectives of the test items. In Pakistan generally, four options are used for one stem. These options contain correct answer called key, and other options are called distractors which are not correct answers but closely related to key or correct answer. These options are designed under rules i.e how these options may be arranged, the options may belong to one category and class etc. These MCQs are used extensively as a formative assessment and summative assessment tool and also for achievement tool in much experimental research. Cizek and O'Day (1994) reported that test Item (MCQ) mostly consists of a stem of question statement and a few possible answers called options. Out of all these options one option represents correct answer called key while other options are called distractors. Now a days MCQs items are used to measure higher order thinking skills such as critical thinking skills, analytical skills and interpretation skills in the framework of Blooms Taxonomy (Kumar, Jaipurkar, Shekhar, Sikri &Srinivas, 2021).

Multiple choice test items are very convenient and easy to attempt for students. A large portion of the science curriculum can be measured in very short period. They help efficiently in identifying the weaknesses and strengths of science students. They provide guidance to the teachers to improve their skills (Tan & McAleer, 2008). Properly constructed items can test HOTS of Blooms taxonomy (Carneson, Delpierre& Masters (2011). Multiple Choice Items are objective in nature and is source of minimizing the researchers/ teachers' biases (Vyas & Supe, 2008). All this

is possible only if the items have processed through the procedures of item analysis.

Difficulty Level/ Index (DIF. I)

Difficulty level/index of test items is essential component of item analysis, defined as the major proportion of students who select the correct options in a test item. If less umber of respondents selects the correct options of key, then this test item is difficult. Difficulty index ranges from 00% to 100%. Difficulty level/index can be found by using this formula when 3rd part of high achiever and 3rd part of low achiever are selected out of the whole sample.

$$P= C/N \times 100$$

Here "P" represents difficulty index, "C" represents number of respondents who attempted right option/answer and "N" represents the total number of respondents/students.

Discrimination Index (DI)

Discrimination index (DI) is a domain of item analysis where it is calculated that how much an item discriminates between high achievers and low achievers. Range of discrimination index is from -1 to +1. Gujjar, Kumar and Rana (2014) define the discrimination index as "the ability of an item to differentiate between students of higher abilities and lower abilities". Discrimination index can be calculated with the help of this formula: $DI= 2 (R_H-R_L) /R_H+R_L$

Discrimination index value may be negative when low achievers are greater in numbers who select the correct answer.

Distractor & Distractor Efficiency (DE)

As discussed in MCQs each question statement has answers in the form of options. All options except correct option are called distractors. Distractors are wrong answers. Distractor analysis is done to assess, whether a student being tested is able to visualize difference in a test. A test developer must follow some rules while developing test items. And distractors are so closely related that it becomes difficult for a respondent to choose correct answer. Some researchers say that a good distractor attracts more respondents with low abilities. According to "Malau-Aduli

and Zimitat (2012)", a distractor which is not attempted by any respondent is dysfunctional as it does not help to measure the educational objectives, is valueless for test item and has negative impact on learner. Mehta and Mokhasi (2014) stressed that distractors are essential component of a test item as it has a reasonable impact on the total test scores. Students' performance is related to the design of the distractors. Distractors on the basis of their functions are categorized into two types. They are dysfunctional distractor (NFD) and functional distractor (FD). This division can be inferred on the basis of respondent's responses. If an option is choiced by less than 5 % respondents (students), it is considered as dysfunctional/ nonfunctional/ ineffective distractor. Conversely if distractor is selected by more than 5% students, it is termed as functional /effective distractor. Muhammad, Tarrant, Ware (2009), Vyas,Supe (2008), & Patil and Patil (2015), suggested that the distractors/ options which are selected by more than 5% students are called functional distractors (FDs) and distractors/options which are selected by less than 5% students then it is called dysfunctional/nonfunctional distractors (DFDs/NFDs). Distractors can be represented in terms of percentage by this formula:

Distractors Percentage = Number of students who selected distractor/total number of students x100

Effectiveness of distractor or its efficiency ranges from (0-100) % and is determined on the basis of functional or dysfunctional distractors in an atom. Distractor Efficiency (DE) of an item with one key and three distractors can be expressed as 100%, 66.66%. 33.33% and 0% depending upon the number of dysfunctional distractors (DFD)

Previous Research

Research study conducted by Sharma (2021) on item analysis on B.Ed students in Nepal in 2020 on 27 students with 20 Multiple choice questions, difficulty level (Diff I), discrimination index (DI) and distractor analysis (DE). Three (3) items had Diff.I level between 0.20- 0.29, fourteen (14) items between (0.40-0.59), two (2) items between (0.60-0.79) and one (1) item had (0.80-0.89). The discrimination index of two 920 items was between (0.20-0.29), there was no item with discrimination index (0.30-0.39) and Eighteen (18) was greater than 0.40.Sharma found that there

were five (5) items with one dysfunctional distractor, fifteen (15) items with zero dysfunctional distractor.

Another research study conducted by (Agarwal, Burud and Nagandla, 2019) on 120 multiples choice questions. Sample for data collection was 113 students of "International Medical University, Malaysia" This study found that there were five (5) items were very difficult, twenty (20) items were good, forty-two (42) items excellent, and eighteen items easy and thirty-five (35) items were very easy. The Discriminatory efficiency found in this research was that there were forty-five (45) items with good discrimination, thirty-seven (37) with fair discrimination, twenty-three (23) showed poor discrimination and fifteen (15 items showed negative discrimination efficiency. In the same way there were forty-seven (470 items with no dysfunctional distractors, fifty-one items (51) with one dysfunctional distractor, eighteen (18) with two dysfunctional distractors and four 94) items with 3 dysfunctional distractors.

In Item analysis research conducted byAlam, Butt, Hassan,Konain, Mahjabeen and and Rizvi (2018) on 65 test items (MCQs). Sample for this research was 110 students studying in 4th year MBBS programme at "Islamabad Medical and Dental College Islamabad" during the year 2017. The findings of this research study regarding difficulty level were varying. There was only one item very difficult, fifty-three (53) items difficult but acceptable and eleven (11) items were too easy. There were thirty-four items with excellent discrimination, fifteen items showed good discrimination, five items represented acceptable discrimination and eleven (11) items depicted poor discrimination index. It was found in this research that there were sixteen (16) items with zero dysfunctional distractors, thirty (30) items had one dysfunctional distractor, sixteen (16) with two DFD and three (3) items were with three dysfunctional distractors.

Research conducted by Boratne, Palve, Patil and Vell(2016) conducted a study with 30 MCQs on 22 students at research institute at Mahatma Gandhi Medical College. Total 30 keys and 90 distractors were analyzed. Average of difficulty level (Diff. I), discrimination index (DI) and distractor efficiency (DE) was 38.3%, 0.27 and 82.8% respectively. Eleven test items out of 30 showed higher difficulty index while five (5) showed difficulty easiness greater than 60. (Symbolically Diff. I of 11 test

items >30% &Diff.I of 5items > 60%). Discrimination index of 15 test items was very good. It was amazing that 16 distractors with percentage of 17.8% were found dysfunctional being selected by less than 5% respondents. Research on item analysis by Namdeo and Sahoo (2016) conducted on 25 MCQs. Data was collected from 76 students of medical at Kalinga Institute of Medical Science (KIMS) Bhubaneswar showed that Diff I of 8 test items with diff I > 70% were too easy, 14 items with diff I range (30-70%) were acceptable whereas 3 test items with diff I < 30% were too difficult. Discrimination index (D I) value of 12 test items >0.35 was excellent, range of DI of 3 test items (020- 0.34) and 8 test items was found < 0.2. There was total 75 distractors, there were 22 test items with dysfunctional distractors, whereas 8 test items contained one dysfunctional distractor, 10 test items have two dysfunctional distractors and 4 test items contained 2 dysfunctional distractors (Namdeo& Rout, 2016).

Methodology

Design of the Research

Cross-sectional survey research design which is a type of descriptive research was used for this study. Researcher gathered primary data from the test which was used in research to find out the effect of e-modules on the academic performance of 7th class general science students. Researcher analyzed the data to do item analyses and to find the difficulty level (Diff I), discrimination index (DI) and distractor efficiency (DE).

Sample and Sampling Technique

Sample of two hundred and forty (240) students studying in 7th class in Islamabad Model Institutions under FDE, Islamabad from all students studying in 7th class during academic year 2022-2023. The sample was selected through multistage random sampling technique. First two sectors one from out of six sectors were randomly selected. In Second stage one school from each selected sector was randomly selected. In last stage 120 students studying in 7^{th} class were randomly selected from each school. In this way test was administrated on 240 students.

Variables

Researcher developed 30 test items with four options (one correct options called key and three very close to correct option but wrong options called distractors in each test item) are major variables. Difficulty level (Diff I), Discrimination Index (DI) and distractor Efficiency (DE) are other variables of this study.

Data Tool and Source of the Data

The primary source of data was test which was developed from four units of general science (Chemistry Section) from class 7th curriculum for students studying in schools of Islamabad Capital territory. Data was ratio data. Test as a tool of collecting data was developed in the form of MCQs. Test was primarily validated by the experts and academicians.

Data Analyses & Results

All data were analyzed. Test was administered to 240 students, marked according to key. Tool contained 30 test items. Frequencies and percentages were used to calculate the item analysis. In first step marks of the students were arranged in ascending order and 240 students were split into three parts. Each part contained 80 students. Eighty students who got more marks and 80 students who got least marks were selected.

For analysis and evaluation of difficulty index/level of the test items researcher took help from the following table. Test items of general science (Chemistry section) from subarea of Chemistry were evaluated as "1" for correct answer and "0" for wrong answer and difficulty level (P) was calculated as per formula. Thirty test items were developed in the chemistry. This tool was presented to experts for validation. As a result of expert opinion test item no.1, 24 and 27 were eliminated due to ambiguity and repetition. Item number 24 was repeated. Remaining tool with 27 test items were administered to 240 students of class 7^{th}. Marks were awarded to each student against each test item. Difficulty level was calculated for results. Detail is given in table 1

Table.1

Difficulty level (P) of 30-3=27 Items in the area of Chemistry

S.No	C	N	P= C/N x100	P	S.No	C	N	P= C/N x100	P
1/17	66	240	66/240	27.75%	15/14	115	239	115/239	48.12%
2/18	64	239	64/240	26.78%	16/16	114	238	114/238	47.90%
3/5	92	240	92/240	38.33%	17/19	141	240	141/240	58.75%
4/8	90	238	90/238	37.78%	18/20	140	240	140/240	58.33%
5/9	73	240	73/240	30.41%	19/21	140	239	140/239	58.58%
6/15	72	237	72/237	30.37%	20/22	139	237	139/237	58.64%
7/23	93	240	93/240	38.75%	21/25	133	238	133/238	55.88%
8/2	119	239	119/239	49.79%	22/26	119	240	119/240	49.58%
9/6	122	240	122/240	50.83%	23/28	126	240	126/240	52.25%
10/7	134	238	134/238	56.30%	24/29	122	240	122/240	50.83%
11/10	137	240	137/240	57.08%	25/30	126	239	126/239	52.74%
12/11	121	239	121/239	50.62%	26/3	190	240	190/240	79.17%
13/12	120	240	120/240	50,00%	27/4	185	240	185/240	77.08%
14/13	118	239	118/239	49.37%		Total Sum of 1 to 27			1410.07
Average Difficulty Level of 27 test items									52.22
Total Sum after removing item no 17,18 (too much difficult)									1355.54
Average Difficulty Level after exclusion of item no.17,18									54.22

Table 2

Evaluation and Recommendation about difficulty index (Diff I)

S.N.	Difficulty Index	N	Item Evaluation	Recommendation
1.	>0.30	2	Very hard/ Most difficult	Eliminated
2.	0.30-0.39	5	Difficult	Revised and improved
3.	0.40-0.59	18	Moderately difficult	Kept and sustained
4	0.60-0.79	02	Easy	Kept
5	.080-0.89	00	Too easy	---------

From the table 1, item numbers 1, 24 & 27 were removed as a result of first review of experts. Table.2 represents that there are two test items 17 & 18 with very high difficulty level. There are five (5) test items 5, 8, 9, 15 & 23 were difficult and were revised. There were 18 test items viz 2, 6, 7, 10,11,12,13, 14, 16, 19, 20, 21, 22, 25, 26, 28 & 30are which were

moderately difficult and sustained. Only two test items 3 & 4 were found easy. There was no item found to be very easy.

Table.3

Discrimination Index (DI) for the tool (Test of Chemistry)

S.N	R_U	R_L	Formula=$2(R_U-R_L)/R_U+R_L$	D I	Interpretation
1	52	27	2(52-27)/80+79	0.31	Discriminating items, good items
2	65	60	2(65-60)/80+80	0.06	Marginalized test items with minor discrimination
3	70	49	2(70-49)/80+80	0.26	Moderately discriminating, fair items
4	46	13	2(46-13)/80+80	0.41	Very discriminating, very good item
5	65	12	2(65-12)/80+80	0.66	Very discriminating, very good item
6	64	22	2(64-22)/80+80	0.53	Very discriminating, very good item
7	49	8	2(49-8)/79+80	0.52	Very discriminating, very good item
8	31	14	2(31-14)/80+80	0.21	Moderately discriminating, fair items
9	66	22	2(66-22)/80+80	0.55	Very discriminating, very good item
10	59	19	2(59-19)/80+80	0.50	Very discriminating, very good item
11	53	19	2(53-19)/80+39	0.43	Very discriminating, very good item
12	50	19	2(50-19)/80+80	0.39	Discriminating items, good items
13	52	21	2(52-21)/80+79	0.39	Discriminating items, good items
14	39	5	2(39-5)/79+80	0.43	Very discriminating, very good item
15	54	21	2(54-21)/79+79	0.42	Very discriminating, very good item
16	38	6	2(38-6)/79+79	0.41	Very discriminating, very good item
17	28	13	2(28-13)/80+80	0.19	Marginalized test items with minor discrimination
18	57	39	2(57-39)/79+80	0.23	Moderately discriminating, fair items
19	58	35	2(58-35)/80+80	0.29	Moderately discriminating, fair items
20	58	29	2(58-29)/80+80	0.36	Discriminating items, good items
21	54	41	2(54-41)/80+80	0.16	Marginalized test items with minor discrimination
22	33	23	2(33-23)/79+79	0.13	Marginalized test items with minor discrimination
23	50	37	2(50-37/80+80	0.16	Marginalized test items with minor discrimination
24	52	27	2(52-27)/80+80	0.31	Discriminating items, good items
25	61	24	2(61-24)/79+80	0.47	Very discriminating, very good item

26	60	19	2(60-19)/80+80	0.51	Very discriminating, very good item
27	61	24	2(61-24)/79+79	0.47	Very discriminating, very good item

Table 4

Evaluation and recommendation of discrimination index (Chemistry)

S.N.	DI	N	%age	Evaluation	Recommendations
1.	Negative	0	0	Worst/ defective item	Discard
2.	> 0. 20	05	18.52	Not discriminating item, marginal item	Revised / Discarded
3.	0.20- 0.29	04	14.81	Moderately discriminating, fair item	Kept/Sustained
4.	0.30- 0.39	05	18.52	Discriminating items, good items	Kept/Sustained
5.	≥ 0.40	13	48.15	Very discriminating, very good item	Kept/Sustained

Table.4 depicts that there are five items which showed no discriminatory index with less than 20% DI. This means they were too hard and needed extensive revisions. All were revised. There were four test items which were moderately discriminating and were revised also. All other 18 items were retained out of which five items had discriminatory value between 0.30-0.39 and other thirteen items had discriminatory index value greater than 0.40.

Table5

Distractor Analysis from three distractors and one answer (Chemistry Section)

S.N	N (R.R	Option A	Option B	Distractor C	Distractor D	Correct
½	239	119(49.79%)	55 (23.01%)	51 (21.34%)	*14 (5.85)*	A
2/3	240	13 (5.41%)	190 (79.17%)	17 (07.08%)	20 (8.33%)	B
¾	240	185(77.08%)	30 (12.5%)	13(05.41%)	12(5.0%)	A
4/5	240	81(33.75%)	45(18.75%)	92 (38.33%)	22(9.17%)	C
5/6	240	72(30.00%)	20(8.33%)	26(10.83%)	122(50.83%)	D
6/7	238	65(27.31%)	134 (56.30%)	19(07.98%)	20(8.40%)	B
7/8	238	85(35.71%)	41(17.23%)	90 (37.82)	22(9.24%)	C

8/9	240	70(29.17%)	64(26.67%)	73 (30.43%)	33(13.75%)	C
9/10	240	57(23.75%)	27(11.25%)	137 (57.08%)	19(07.92%))	C
10/11	239	35(16.64%)	121 (50.62%)	52 (21.76%)	31(12.97%)	B
11/12	240	39(16.25%)	120(50.00%)	3715.42%	44(18.33%)	B
12/13	239	75(31.38%)	25 (10.46%	21(07.79%)	118 (49.37%)	D
13/14	239	83(34.72%)	115(48.12%)	20 (8.36%	22(9.20%)	B
14/15	237	67(28.15%)	62(26.05%)	72 (30.38%)	37(15.55%)	C
15/16	238	46(19.33%)	114 (47.90%)	54(22.69%)	26(10.92%)	B
16/17	240	92 (38.33%)	72 (30.25%)	*10 (04.20%)*	66 (27.50%)	D
17/18	239	78(32.64%)	64 (26.78%)	86 (35.98%)	*11(4.6%)*	B
18/19	240	141(58.75%)	41(17.08%)	32(13.33)	26(10.83%)	A
19/20	240	140(58.33%)	51(21.25%)	25(10.42%)	24(10.00%)	A
20/21	239	47(19.66%)	140 (58.58%)	31(12.97%)	21(8.79%)	B
21/22	237	17(07.17%)	53(22.36%)	28(11.82%)	139 (58.65%)	D
22/23	240	93(38.75%)	57(23.75%))	51(21.25%)	39(16.25%)	A
23/25	238	19(07.98%)	21(8.82%)	65 (27.31%)	133 (55.88%)	D
24/26	240	30(12.50%)	3715.42%)	54(22.50%)	119 (49.58%)	D
25/28	240	52(21.67%)	126 (52.50%)	31(12.92%)	31(12.92%)	B
26/29	240	122(50.83%)	67(27.92%)	23(09.58%)	28(11.67%)	A
27/30	239	23(09.62%)	27(11.30%)	63(26.36%)	126 (52.72%)	D

Primarily thirty (30) test items were developed from chemistry section of class 7th general science. First, this test (tool) was validated from three experts. After this process three test items were removed from the tool and five test items were improved. After this item analysis was conducted for remaining twenty-seven (27) test items. There were three distractors for each item and one answer for each item. Total numbers of distractors were eighty-one (81) in this tool along-with twenty-seven correct answers. It was found that there were two distractors with less than 5% distractor efficiency. One was distractor "C" in item number 16/17 and second was distractor "D" in item number 17/18. These two test items were removed as they had high difficulty level and distractors with less than 5% DE were not possible to be replaced with any other distractors hence removed from the main tool.According to Hingorjo& Jaleel (2012) the

distractor selected by respondents (students) less than 5% has very low effectiveness, hence may be revised or removed because it affects the quality of overall test item. Qamar, Kanwal & Nadeem (2022) recommended that distractors with less than 7% efficiency may be considered for revision. So, in this analytical study distractors with less than efficiency was revised and improved. There were eleven distractors with less than 10% efficiency.

Distractor "A" & "D" in item number 2/3, distractor "D" of item 3/4, distractor "D" of item 4/5, distractor "B" of item number 5/6, distractor "C" & "D" of item 6/7, distractor" C" of item 13/14, distractor "D" of item 20/21, and distractor "A" of item 23/24 were dysfunctional. These eleven distractors were revised and improved. Further items numbers 3/4, 4/5, 5/6, 13/14, 16/17, 17/18 & 20/21 had one distractor. But item numbers 2/3 & 6/7 had two distractors with less than 10% distractor efficiency. Overall, thirteen distractors were so called dysfunctional out of which eleven were improved but two were eliminated.

Discussions and Conclusion

One of the techniques in ensuring quality of education is through measurement and evaluation. In the same way there are many techniques and processes for quality assessment and one of them is item analysis. MCQs test items is the best tool for measurement of cognitive abilities and performance of the students (Hingorjo& Jaleel, 2012). The tools and techniques to find out the characteristics and quality of test items frequently used are difficulty index (Diff.I), discrimination index (DI) and distractor efficiency (DE). In this study the most difficult test items have difficulty level less than 30% and were only two test items. In the same way revision of those test items was done which had DI value >0.20. there were five test items whose DI value was less than 0.20. Distractor analysis is done to ensure the quality of distractor. The tool which had maximum functional distractors are acceptable to be accurate for measurement and assessment. According to Hingorjo& Jaleel (2012) the distractor selected by respondents (students) less than 5% has very low effectiveness, hence may be revised or removed because it affects the quality of overall test item. Qamar, Kanwal & Nadeem (2022) recommended that distractors with less than 7% efficiency may be considered for revision. In this study

all those items were revised and improved which had distractor efficiency less than 10%. It is concluded that out of thirty test items five test items were totally removed, eighteen test items were retained but seven test items are revised. Finally, twenty-five test items were used for the quality results.

References

Atalmiş E, Kingston N.(2017) Three, four, and none of the above options in multiple-choice items. *Turkish Journal of Education*. 2017; 10;6(4):143-57. (doi. org/10.19128/turje.333687.)

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.

Carneson, J., Delpierre, G., & Masters, K. (2011). *Designing and Managing MCQs: Appendix C: MCQs and Bloom's Taxonomy*; 2011.

Case, S. M., & Swanson, D. B. (1998). *Constructing written test questions for the basic and clinical sciences* (2nd ed., pp. 22-25). Philadelphia: National Board of Medical Examiners.

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.

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

Elgadal, A. H., &Mariod, A. A. (2021). Item Analysis of Multiple-choice Questions (MCQs): Assessment Tool For Quality Assurance Measures. *Sudan Journal of Medical Sciences, 16*(3), 334-346.

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

Hingorjo, M. R., & Jaleel, F. (2012). Analysis of one-best MCQs: the difficulty index, discrimination index and distractor efficiency. *JPMA-Journal of the Pakistan Medical Association*, *62*(2), 142.

Ismail, F. K. M., &Zubairi, A. M. B. (2022). Item Analysis of a Reading Test in Sri Lankan Context Using Classical Test

Theory. *International Journal of Learning, Teaching and Educational Research, 21*(3).

Jannah, R., Hidayat, D. N., Husna, N., &Khasbani, I. (2021). An item analysis on multiple-choice questions: a case of a junior high school English try-out test in Indonesia. *Leksika: Jurnal Bahasa, Sastra dan Pengajarannya, 15(1),* 9-17.

Kumar, D., Jaipurkar, R., Shekhar, A., Sikri, G., & Srinivas, V. (2021). Item analysis of multiple-choice questions: A quality assurance test for an assessment tool. *medical journal armed forces india*, 77, S85-S89.

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.

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.

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.

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.

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.

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

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.

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

Qamar, A. M., Kanwal, W., & Nadeem, H. A. (2022). Item analysis for test to examine the effect of e-module on the academic performance of 7th class science students in Islamabad. *Jahan-e-Tahqeeq*, *5*(2), 188-196.

Shad, M. N., Fatima, A., Fatima, S. H. A. H. N. A. Z., &Chiragh, S. (2018). Item analysis of MCQs of a pharmacology term exam in a Private Medical College of Pakistan. *Pakistan Journal of Medical & Health Sciences*, *1*(12), 2.

Sharma, L. R. (2021). Analysis of difficulty index, discrimination index and distractor efficiency of multiple-choice questions of speech sounds of English. *International Research Journal of MMC, 2(1)*, 15-28

Sowdani, K. M., Jaber, H., & Thomas, A. M. (2018). Item Analysis as a Tool for Educational Assessment as Compared to Students, Evaluation to lectures. *Al Mustansiriyah Journal of Pharmaceutical Sciences*, *18*(2), 105-113.

Tan, L. T., & McAleer, J. J. A. (2008). The introduction of single best answer questions as a test of knowledge in the final examination for the fellowship of the Royal College of Radiologists in Clinical Oncology. *Clinical Oncology*, *20*(8), 571-576.

Toksöz, S., &Ertunç, A. (2017). Item analysis of a multiple-choice exam. *Advances in Language and Literary Studies*, *8*(6), 141-146.

Vyas, R., & Supe, A. (2008). Multiple choice questions: a literature review on the optimal number of options. *Natl Med J India*, *21*(3), 130-3.