

The Relationship Between Iraqi EFL Teachers' Creativity and Designing Multiple-Choice Questions: A Case Study of Najaf

Ghabanchi, Z. (Ferdowsi University of Mashhad)

ghabanchi@um.ac.ir

Al-Khafaji, H. (Ferdowsi University of Mashhad)

Hussein8794@gmail.com

Abstract

This study aimed to explore the relationship between teachers' creativity and their ability in designing Multiple Choice Questions. Moreover, it attempted to investigate the relationship between gender and the amount of flaws in designing Multiple-Choice Questions (MCQs). To this end, 100 EFL Iraqi teachers who were teaching English at high schools of Najaf were randomly selected. They included 50 male and 50 female. Creativity questionnaire was given to the EFL teachers to rate their creativity. Then, they were asked to design at least five (MCQs) about specific passage to discover is there any flaw in designing MCQs, if yes what is the amount of each group (males/females) and the difference between them. After gathering the data, the researchers calculated the creativity of each person and recorded all required information in SPSS software. Kolmogorov-Smirnov test was used to test the assumption of the normality of the variables. Pearson Correlation was used to test the correlations between the variables. The results indicated that there was no significant relationship between the variables. Moreover, results portrayed that there was no significant relationship between Iraqi EFL teachers' gender and the type of flaws in designing MCQs. The findings also demonstrated that the role of teacher's creativity has no association with their ability in design flawless MC items. The findings concluded that there were several problems in the designed MCQs by Iraqi teachers.

Keywords: *Creativity; design; multiple-choice questions*

1. Introduction

Assessing students' performance in their academic affairs seems very important in all fields. Testing students and teaching are so inter-related that it is virtually impossible to work in either field without constantly concerned with the other (Heaton & Harmer, 1975). Designing a test should not be considered as an easy act because a test first is a method, it has specific instruments, techniques and procedures (Brown & Abeywickrama, 2010). Tests are designed in different methods but specially during last decades procedure of designing tests changed (Madsen, 1983) "testing during last century and the early decades of this one was basically intuitive, or subjective and depend on the personal impression of teachers. ... Subjective tests began to be replaced by objective tests because the latter could be scored consistently even by untrained people" (pp. 5-6).

The most popular type of objective tests is Multiple Choice Questions (MCQ) (Farhady, Ja'farpur, & Birjandi, 2007). MCQs are today very popular and use in different aspects of language learning and this popularity has some reasons; high reliability, convenience in scoring, efficiency, and economy (Ghafournia, 2013). As each MCQ seems consisted of a few lines, many teachers believe they are easy to construct but they are actually very difficult to

The Relationship Between Iraqi EFL Teachers' Creativity and Designing Multiple-Choice Questions: A Case Study of Najaf

Ghabanchi, Z. & Al-Khafaji, H

design correctly (Brown & Abeywickrama, 2010). In recent decades, its advantages are regarded so highly and it seems that it is the only method and way of test. The frequent use of it is in testing of large number of candidates (Hughes, 2003). As mention, MCQs are one of the most difficult items to write, so it is reasonable that they become time consuming too.

Inappropriateness of MCQs seems to come from two main sources, first one is content, and the other is technique (Hughes, 2003). MCQs are invested in many fields, consequently the current paper intends to investigate the relationship between teachers' creativity and their ability in designing Multiple Choice Questions. As for as the significant relationship between Iraqi EFL teachers' gender and the amount of their flaws in designing MCQs.

2. Purpose of the Study

Due to the importance of Multiple-Choice Question (MCQ) in testing students and teaching, this study aims to shed some light on the relationship between Iraqi EFL teacher's creativity and their ability in designing MCQs.

To be more exact, the present research sets out to answer the following questions:

- 1- Is there any significant relationship between designing flawless multiple-choice questions and Iraqi EFL teachers' creativity?
- 2- Is there any significant relationship between Iraqi EFL teacher's gender and the amount of their flaws in designing MCQs?

2.1. Research Hypotheses

H01: There is no significant relationship between designing flawless multiple-choice questions and Iraqi EFL male and female teachers' creativity.

H02: There is no significant relationship between Iraqi EFL teachers' gender and the type of flaws in designing MCQs.

3. Theoretical Background

3.1. Multiple Choice Questions

Textbooks include many MCQs and they are used on high-profile English language proficiency (Coombe, 2010). Any standard MCQs has two basic parts: a problem (the stem) and a list of suggested solutions (alternatives) (Hotiu, 2006). A stem is developed to provide the test takers with "a complete idea of the problem to be solved in selecting the right answer" (Haladyna, 1994, p. 35). Since stems supply the problem under question they are alternatively referred to as "problems" or "leads" as well (Marshall and Hales, 1971; Payne, 1992). While a list of suggested solutions *alternatives*, *alternates*, *choices* or *options*. The correct, best or most appropriate solution is referred to as the *answer*, the *keyed answer*, or *keyed response*, whereas the remaining options are called *distracters* or *foils* (Khodadady, 1997).

Unfortunately many teachers believe that MCQs are superficial exercises and do not need much understanding and thought (Clegg & Cashin, 1986). Although they seem to be very simple, they are extremely difficult to design correctly (Brown & Abeywickrama, 2010). There are many reasons that show why these kinds of tests are widely used among teachers (Simkin & Kuechler, 2005). They result in more content valid interpretation for test score, reliability of these tests seems very high and they are easy to use and store (Haladyna & Downing, 1989). Other researchers (Clegg & Cashin, 1986) and (Hotiu, 2006) claimed that MCQs could be used for evaluating higher level of learning, e.g. contrasting and comparing. (Clegg & Cashin, 1986) also mentioned that MCQs are used in university classes widely and emphasized that MCQs can be used as diagnostic purpose for understanding strengths and

weaknesses. Teachers can use them effectively for measuring students' understanding of class materials (Simkin & Kuechler, 2005).

Grammatical inconsistency between the stem and the response, providing cues that help students to choose correct response, and specific words that show gender bias are among common types of violations that may occur in a test (Coombe et. al., 2010).

3.1.1.Guidelines for Designing Multiple Choice Questions

There are an essential procedures for designing MCQs that teachers should try to follow them as best as possible when writing multiple choice test questions.

- There should be 3 to 4 answer options.

Items should have enough answer options that pure guessing is difficult, but not so many that the distractors are not plausible or the item takes too long to complete.

- All answer options should be plausible.

If an answer option is clearly not correct because it does not seem related to the other answer options or is from a content area not covered by the test, it does not operate as a distractor.

- Distractors should be clearly incorrect.

Always double check to make sure that correct answer is fully correct and that distractors are definitely wrong (and that there isn't room for someone to argue that it is indeed accurate).

- "All of the Above" and "None of the Above" should not be answer options.

Some students will guess this answer option frequently as part of a test-taking strategy. Other students will avoid it as part of a test-taking strategy. Either way, it does not operate fairly as a distractor. Measuring this particular analytic ability is likely not the targeted goal of the test.

- Answer options should all be grammatically consistent with stem.

If the grammar used in the stem makes it clear that the right answer is a female or is plural, teachers make sure that all answer options are female or plural.

- Answer options should not be longer than the stem (and equally long with each other).

An item goes more quickly if the bulk of the reading is in the stem, followed by brief answer options. Teachers should try to make the response option shorter than the stem and about equal in length.

- Order of answer options should be logical or random.

For obvious reasons, the answer should not always be "C". This can be done by placing the answer options in an order based on some rule (e.g. shortest to longest, alphabetical, chronological) or by randomizing the order of answer options once the test is constructed.

- Negative wording should not be used.

Some students read more carefully or process words more accurately than others, and the word "not" can easily be missed. Even if the word is emphasized so no one can miss it, educational content is more likely stored as a collection of positively worded information.

The Relationship Between Iraqi EFL Teachers' Creativity and Designing Multiple-Choice Questions: A Case Study of Najaf

Ghabanchi, Z. & Al-Khafaji, H

- Stems should be complete sentences.

If a stem is a complete question, ending with a question mark, or a complete instruction, ending with a period, students can begin to identify the answer before examining answer options. Students must work harder if stems end with a blank or a colon or simply as an uncompleted sentence. More processing increases chances of errors (Haladyna & Downing, 1989a, 1989b).

3.2. Creativity

Many researchers interested in investigating the relationship between creativity and teachers' specific features (Vasudevan, 2013), (Reber, 1995), (Pishghadam, Ghorbani Nejad, & Shayesteh, 2012), (Ghafourian, 2012), (Ghonsooly & Raeesi, 2012). (Blumen-Pardo et al., 2002; Crowther et al., 2008) claimed, "Creativity normally helps teachers resolve many unexpected incidents and assist them build a better teaching environment" (cited in (Maktabi, Hanifi, & Feizabadi, 2013)) (P. 1). Many debates are around the definition of creativity, but briefly, "creativity is the production of novel and appropriate ideas or works." (Baghaei & Riasati, 2013) (P. 1)

Some researchers (Kandi & Kandi, 2013) emphasized importance of creativity by stating, "The most complex and outstanding perspectives of human thought is creative thinking, about which different views exist. The human's creativity is the most important equipment by which he can eradicate spiritual pressures of daily life and extraordinary events." (P. 1)

It was stated (Pishghadam, et al., 2012) that "In today's modern world, barely anyone interrogates the role of creativity. To a society that constantly needs to prosper and survive, creativity is an inevitable utensil. Unlike abundant concepts in science, there is no unified, unambiguous definition in terms of this mysterious notion." (P. 1)

4. Methodology

4.1. Participants

For this study, among many EFL teachers in Najaf 100 EFL teachers who are teaching English in high school of Najaf are selected. Selection of participants is based on Krejcie & Morgan (1970) Sample Size Table with 95% confidence and 0.05 degree of accuracy. EFL teachers are chosen but only 100 of them participate in this study. These samples are teaching English at public school of Najaf in 2016-2017 academic years in different areas. The participants include 50 male English teachers and 50 female English teachers. All teachers are at B.A level in one field of English Language and most of them are familiar with MCQs flaws that may happen in a test.

4.2. Instrument

As this study investigates the relationship between teachers' creativity and their skill in designing flawless multiple-choice questions, two instruments are used:

1. Abedi (1985) creativity questionnaire is given to the EFL teachers. The reliability and validity coefficient of test of creativity has been estimated to be %90 to %72. Each person should read its items carefully and answer to items based on provided instruction. It includes 60 items and the ranges of scores are between zero to 120. It takes about 15 minutes to answer to the included items.
2. A certain passage will be given to the teachers. Researcher will distribute twenty copies to the EFL teachers male and female in equal then he asks them to design At least five MCQs about this passage.

The Relationship Between Iraqi EFL Teachers' Creativity and Designing Multiple-Choice Questions: A Case Study of Najaf

Ghabanchi, Z. & Al-Khafaji, H

4.3. Procedure

For collecting data, researcher distributes the Creativity Questionnaire to EFL teachers. All participants are EFL teachers and researcher assures them that their gathered information keep confidential. Teachers' selection is done randomly among Najaf teachers who have B.A. degree. For this reason, researcher sends require questionnaire and MCQs test to EFL teachers' email and gives to some of them printed paper format. After gathering data, researcher will calculate the creativity of each person and records them.

After distributing the designing MCQs test on EFL teachers equally, Ten male and ten female. Researcher asks them to design At least five MCQs about the passage that mentioned before. Researcher gathers teacher's answers then he records if there any significant relationship between Iraqi EFL teachers' gender and the amount of their flaws in designing MCQs.

After gathering all required information, researcher records them in SPSS software (Version 21) for gaining statistical information. Then researcher correlates teachers' creativity in designing MCQs to see whether his hypothesis is confirmed or rejected.

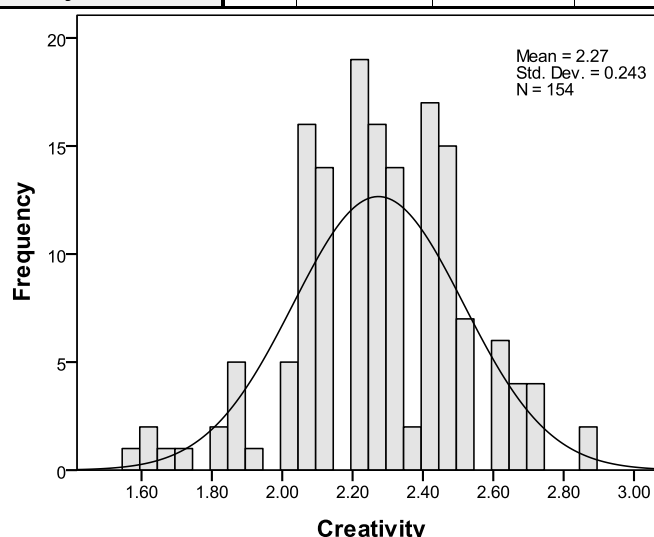
5. Results

The first question of the study is whether there is any significant relationship between designing flawless multiple-choice questions and Iraqi EFL teachers' creativity.

Descriptive statistics are reported to summarize the characteristics of the data including minimum, maximum, mean, standard deviation values of creativity. As can be seen, the lowest and highest creativity is respectively 1 and 3 and also the mean and standard deviation are respectively 2.27 and 0.24.

Table5. 1: *Descriptive Statistics for creativity*

variables	N	Minimum	Maximum	Mean	Std. Deviation
creativity	154	1	3	2.27	0.24



The following table shows that 8.3% of Iraqi teachers suffer from the Grammatical clue flaw, 13.3% of Grammatically incorrect items flaw, 11.7% of Grammatical inconsistency flaw, 18.3% of Unparalleled options flaw, 13.3% of flaw Unreasonable options, 11.7% of the

The Relationship Between Iraqi EFL Teachers' Creativity and Designing Multiple-Choice Questions: A Case Study of Najaf

Ghabanchi, Z. & Al-Khafaji, H

Double answer key flaw, 3.3% of the flaw of the None of the above, 0.5% of the Negative Stem flaw, 10.10% of the Word repeat flaw, 7.1% of the Start with blank flaw and 3/3 of Impure items flaw. By comparing the results, it can be said that unparalleled options flaw is the most common one in the designing of multi-choice questions among Iraqi teachers.

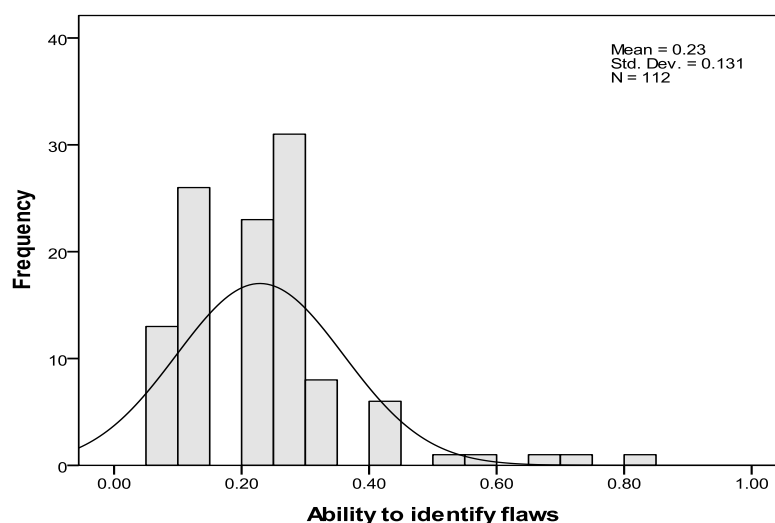
Table5. 2: *Descriptive Statistics for Types of flaw in Designing MCQs*

Types of flaw	Frequency	Percent
Grammatical clue	5	8.3
Grammatically incorrect items	8	13.3
Grammatical inconsistency	7	11.7
Unparallel options	11	18.3
Unreasonable options	8	13.3
Double answer key	7	11.7
Non of the above	2	3.3
Negative stem	3	5.0
Word repeat	6	10.0
Start with blank	1	1.7
Impure items	2	3.3
Total	60	100.0

As table three shows, descriptive statistics are reported to summarize the characteristics of the data including minimum, maximum, mean, standard deviation values. As can be seen, the lowest and highest ability to identify flaws are respectively 0.07 and 0.8, and the mean and standard deviation are respectively 0.23 and 0.13.

Table5. 3: *Ability to Identify Flaws*

variables	N	Minimum	Maximum	Mean	Std. Deviation
Ability to identify flaws	112	0.07	0.8	0.23	0.13



The Relationship Between Iraqi EFL Teachers' Creativity and Designing Multiple-Choice Questions: A Case Study of Najaf

Ghabanchi, Z. & Al-Khafaji, H

To implement statistical methods and calculate the appropriate test statistic and logical deduction about research hypotheses, the most important step before any action is to select the appropriate statistical method for the research. For this purpose, knowledge of the distribution of data is a priority. In this study, Kolmogorov-Smirnov test is used to test the assumption of the normal variables of the research. The statistical constraints of Kolmogorov-Smirnov's normalization test are as follows.

H_0 :data are normally distributed.

H_1 :data are not normally distributed.

Table5. 4: *One-Sample Kolmogorov-Smirnov Test*

	creativity	multiple-choice questions	Identify Flaws
N	154	20	112
Kolmogorov-Smirnov Z	.524	.701	.947
Asymp. Sig. (2-tailed)	.301	.339	.208

Inferential Statistics and Hypothesis Testing

In this section, the hypotheses are examined. According to the normality of the variables, the parametric Pearson correlation tests (Pearson Correlation Analysis), independent t-test (Independent Sample T Test) and chi-square test (Chi-Square Tests) are used to test the hypotheses.

The First Hypothesis

In this hypothesis, the relationship between designing of flawless multi-choice questions and Iraqi EFL male and female teachers' creativity is discussed. Regarding the normality of variables, Pearson correlation test is used to test this hypothesis. The test hypothesis is as follows:

H_0 :There is no significant relationship between designing flawless multiple-choice questions and Iraqi EFL male and female teachers' creativity.

H_1 :There is a significant relationship between designing flawless multiple-choice questions and Iraqi EFL male and female teachers' creativity.

Table 5. 5: *Correlations between designing flawless multiple-choice questions and creativity*

			MCQs
Creativity	Male	Pearson Correlation	.061
		Sig. (2-tailed)	.619
		N	77
	Female	Pearson Correlation	.030
		Sig. (2-tailed)	.817
		N	77

As table five indicates, designing of flawless multi-choice questions is not significantly correlated with Iraqi EFL male teachers' creativity. Pearson correlation coefficient is 0.061. Since p-value is greater than 0.05 in correlation conducted (Sig=.619 >0.05). Accordingly, the

The Relationship Between Iraqi EFL Teachers' Creativity and Designing Multiple-Choice Questions: A Case Study of Najaf
Ghabanchi, Z. & Al-Khafaji, H

null-hypothesis with the 95% confidence is accepted. Based on table five, designing of flawless multi-choice questions is not significantly correlated with Iraqi EFL female teachers' creativity. Pearson correlation coefficient is 0.030. Since p-value is greater than 0.05 in correlation conducted (Sig=.817 >0.05). Therefore, the null hypothesis with the 95% confidence is accepted.

Therefore, it can generally be said that there is no significant relationship between designing flawless multiple-choice questions and Iraqi EFL male and female teachers' creativity and the first hypothesis of the research is confirmed.

The second question of the study is whether there is any significant relationship between Iraqi EFL teachers' gender and the amount of their flaws in designing MCQs. To test the second hypothesis, there is no significant relationship between Iraqi EFL teachers' gender and the type of flaws in designing MCQs; Chi-Square Test is employed.

H₀: There is no significant relationship between Iraqi EFL teachers' gender and the type of flaws in designing MCQs.

H₁: There is significant relationship between Iraqi EFL teachers' gender and the type of flaws in designing MCQs.

As previous results shows that there is no significant relationship between Iraqi EFL teachers' gender and the type of flaws in designing MCQs. The chi-square is 6.604. The significance level of the test (sig) is 0.762, which is more than 0.05 (Sig = .762 > 0.05). Accordingly, the first null hypothesis with the 95% confidence is accepted.

In general, among teachers who have been misunderstood in designing multiple questions, 36.7% are male and 63.3% are female.

Table 5.6: *Type of Gender's Falw*

			Gender		Total
			Male	Female	
Type Of Falw	Grammatical clue	Count	3	2	5
		% within Type of Falws	60.0%	40.0%	100.0%
	Grammatically incorrect items	Count	2	6	8
		% within Type of Falws	25.0%	75.0%	100.0%
	Grammatical inconsistency	Count	3	4	7
		% within Type of Falws	42.9%	57.1%	100.0%
	Unparallel options	Count	3	8	11
		% within Type of Falws	27.3%	72.7%	100.0%
	Unreasonable options	Count	3	5	8
		% within Type of Falws	37.5%	62.5%	100.0%
Type Of Falw	Double answer key	Count	2	5	7
		% within Type of Falws	28.6%	71.4%	100.0%
	Non of the above	Count	1	1	2
		% within Type of Falws	50.0%	50.0%	100.0%
	Negative stem	Count	1	2	3
		% within Type of Falws	33.3%	66.7%	100.0%
	Word repeat	Count	2	4	6
		% within Type of Falws	33.3%	66.7%	100.0%
	Start with blank	Count	0	1	1

The Relationship Between Iraqi EFL Teachers' Creativity and Designing Multiple-Choice Questions: A Case Study of Najaf
Ghabanchi, Z. & Al-Khafaji, H

	% within Type of Falws	.0%	100.0%	100.0%
	Count	2	0	2
Impure items	% within Type of Falws	100.0%	.0%	100.0%
Total	Count	22	38	60
	% within Type of Falws	36.7%	63.3%	100.0%

Table5. 7: *Chi-Square Tests*

Chi-Square	6.604
df	10
Asymp. Sig.	0.762

6. Discussion

The major aims of the present study are: firstly, whether teachers (male and female) with high creativity design a MCQ more accurately and flawless than less creative one. Secondly, to investigate the relationship between Iraqi EFL teachers' gender and the amount of their flaws in designing MCQs.

Teachers' creativity can help students to increase their level of thinking and teachers' communication with students ([Vasudevan, 2013](#)). Although there have not been rare studies aimed at identifying teachers' creativity roles, only some researchers investigate the role of this characteristics on teachers' performance in class activity affairs. Furthermore, the classroom test is one of the most important aspects of the teaching-learning process, and that is a very important responsible for teachers to design their classroom tests ([Clegg & Cashin, 1986](#)). Here in this investigation, the possible relationship between EFL teachers' creativity male and female and designing multiple-choice questions is investigated.

As the results of the study present that there is no significant relationship between designing flawless multiple-choice questions and Iraqi EFL male and female teachers' creativity. In this study, Kolmogorov-Smirnov test is used to test the assumption of the normal variables of the research. Thus; designing of flawless multi-choice questions is not significantly correlated with Iraqi EFL male and female teachers' creativity. This is a reference to Davies's (2006) ideas that teachers indicated a positive attitude towards creativity and the belief that creativity is very important for effective teaching. On the other hand, teachers acknowledged that creativity is not a priority in comparison to such activities as knowledge acquisition, skills training, behavioral management, and motivation development.

The findings also refer that there is no significant relationship between Iraqi EFL teachers' gender and the amount of their flaws in designing MCQs. Cronbach (1970) declared that "the design and construction of achievement test items have been given almost no scholarly attention" (p. 509). Multiple choice items are used in almost all fields of knowledge to measure the mental process of knowledge, comprehension, application, analysis, synthesis and evaluation. Thus; designing MCQs is not depended on gender rather than on knowledge, skills, and practice. So, teachers need more training and courses in writing MCQs to be more familiar with their flaws.

References

Brown, H. D., & Abeywickrama, P. (2010). *Language assessment: Principles and classroom practices* (2 ed.): Longman White Plains^ eNY NY.

The Relationship Between Iraqi EFL Teachers' Creativity and Designing Multiple-Choice Questions: A Case Study of Najaf
Ghabanchi, Z. & Al-Khafaji, H

- Chen, A., Li, L., Li, X., Zhang, J., & Dong, L. (2013). Study on Innovation Capability of College Students Based on Extenics and Theory of Creativity. *Procedia Computer Science*, 17, 1194-1201.
- DiBattista, D., & Kurzawa, L. (2011). Examination of the Quality of Multiple-choice Items on Classroom Tests. *The Canadian Journal for the Scholarship of Teaching and Learning*, 2(2), 4.
- Farhady, D. H., Ja'farpur, D. A., & Birjandi, D. P. (2007). *Testing Language Skills; From Theory to Practice*. Tehran: SAMT.
- Ghafournia, N. (2013). The Relationship between using Multiple-Choice Test-Taking Strategies and General Language Proficiency Levels. *Procedia-Social and Behavioral Sciences*, 70, 90-94.
- Heaton, J. B., & Harmer, J. (1975). *Writing English Language Tests: A practical guide for teachers of English as a second or foreign language*: Longman.
- Hughes, A. (2003). *Testing for language teachers* (2 ed. Vol. 1): Cambridge University Press Cambridge.
- Madsen, H. S. (1983). *Techniques in testing*: ERIC.
- Muhaidib, N. S. A. (2010). Types of Item-Writing Flaws in Multiple Choice Question Pattern – A Comparative Study. 2, 37. Retrieved from <http://uqu.edu.sa> website:
- Sharma, B. (2013). A Study Of Certain Psychological Traits Of Effective And Ineffective Teachers. *Monthly Multidisciplinary Research Journal*, 3(8).