

ISSN 1798-4769

Journal of Language Teaching and Research, Vol. 10, No. 3, pp. 627-640, May 2019

DOI: <http://dx.doi.org/10.17507/jltr.1003.27>

Basis of Designing Test Questions of English Language That Reflect Intended Learning Outcomes (ILOs)

Mona M. Hamad

English Department, College of Science & Arts Muhayil, King Khalid University, Saudi Arabia

Abstract—Teaching is one of the most important professions that affect different aspects of life. Brown (1994) stated that testing and teaching are different, but they are interwoven and interdependent, that it is difficult to tease them apart. Teachers teach, and then test to find out if they had achieved their course objectives. As teachers sometimes fail or find difficulties to come out with perfect test that backwash the whole process of course learning; in other words they fail to construct test that meets all intended learning outcomes (ILOs). So, this paper aims at: finding out reasons behind questions' weakness of English Language tests, and finding solutions that can help educators write their test-questions in scientific pedagogical way, that tests all course content, and reflects (ILOs). The researcher used the descriptive analytical method, a teachers' questionnaire is used as instrument to collect data, and 10 tests' questions were analyzed to tackle reasons behind questions' weakness. Participants of this study are 22 instructors from College of Science & Arts, Muhayil "KKU". Finding indicated reasons behind questions' weakness of English Language tests are the inadequate way of designing tests, as a result of ignorance of the basics of test questions criteria and lacking knowledge of how to use Bloom Taxonomy and (TOS). Instructors need training about how to use Bloom Taxonomy and Table of Specification in designing Tests' Questions of English Language in a pedagogical way.

Index Terms—designing, table of specification (TOS), Intended Learning Outcomes (ILOs)

I. INTRODUCTION

Teaching and testing are integrated, it is impossible to work in either field without being concerned with the other. Scientific measurement ensure accuracy and form the foundation for test theory , Brown (1994) stated that testing and teaching are different, but they are interwoven and interdependent, it is difficult to tease them apart. Teachers teach their students language and then test, they teach them for the whole year then test them to find out if they have achieved their goal or not. Writing test questions is not a matter of opening a book and writing questions that have answers inside the book, it is a scientific process that tests students' learning outcomes, it helps teachers find out if they have achieved their desired planed objectives and learning outcomes of a course or not. During my work as a teacher I used to make my tests the last month before the assigned time for it, and they always test lower level of thinking, with experience, I learnt to start making my test questions from the 1st day I start teaching my course, at the end of the semester I prepare dozens of questions, but not in a helpful way that help me choose the right question for the right learning outcome. Later I learnt how to write test-questions that are aligned with intended learning outcomes, this happened after getting a training titled "How to make a questions' bank" done by the Ministry of General Education in Sudan. After the training, when I was chosen by Sudan Ministry of General Education / Sudan Examinations & Evaluation Directorate in 2009, to feed Sudan Secondary Certificate questions' bank, I practiced what I have learnt. By participating in feeding the questions' bank with 20 exams - questions for S.S.C in 2010, which was a unique and fruitful experience for me that helped me write my test-questions in a pedagogical way.

CBU (2017) pointed that learning outcomes reflect what a person must do or know through the educational process. Learning outcomes are the most important part in curriculum design, as they provide a clear map for designing, delivering and testing procedure

By the end of the any educational process, teachers test students to:

- Evaluate students' performance.
- Identify students' strength and weakness.
- Reinforce learning and teaching.
- See which part of the syllabus needs amendments.
- Gain college credit.
- Provide a way to measure a teacher and/or school's effectiveness.
- Provide a basis for entry into an internship, program, or college.
- Motivate learning when students learn from their mistakes.
- Have a wash back on teaching methods and techniques used by teachers.

Any of the previous learning outcomes or reasons requires testing. As testing is a scientific process. Educator tests samples of what the students learn, so this sample should be chosen carefully to backwash teaching methods and techniques that are applied and what have been learnt.

This paper aims at:

- 1- Finding out reasons behind questions' weakness of English Language tests.
- 2- Finding solutions that can help educators write their test-questions in scientific pedagogical way that test all course content, and reflect intended learning outcomes.

II. LITERATURE REVIEW

Bloom (1965) stated most of our tests test lower level of thinking, then Wiggins G. (1992) pointed to consider higher level of thinking in instruction and assessment, to assess students' ability of problem solving. Wiggins, G. (1992) stated good teaching is inspirable by good testing. McMillan (2003) pointed that variety of strategies employed by teachers in the educational process to teach their students, are reflected by assessment, which illustrate how students experienced during process of learning. McMillan also pointed that Shepard in (1997)(2000a)(2000b) has argued that "traditional assessment perspectives based on traditional behaviorist theories of learning "scientific measurement" and conventional psychometric principle (i.e. those important for large scale standardized objective test) conflict with implication for classroom assessment derived from more contemporary views of learning represented by cognitive constructivist and interpretive paradigms. Brown's (2004) showed the relationship among testing, assessing and teaching, and then showed how teaching overlaps assessment and measurement, he pointed that the reason of testing is to evaluate learners' competence. During ongoing processes, teachers provide instructions along with assessment in order to evaluate how learners acquire and produce what they have learned in the classrooms. Thus, assessment, measurement, tests are interrelated to teaching. Reynolds, Livingston, & Willson (2006) argued that although many teachers love teaching, they have negative interest in testing students. John Biggs and Catherine Tang (2007) stated: *Designing assessment tasks should be aligned to intended learning outcomes ILOs, this assessment task should tell us what and how well our students achieved from the intended learning outcomes, and each assessment task may address several ILOs.* Alexis & Ricardo (2009) highlighted the importance of providing adequate training in language assessment for all prospective language teachers in Colombia, as language teachers are not trained to construct assessment design that helps students to be motivated and engaged or enhanced in the learning process. What Alexis & Ricardo highlighted is probably the same situation in most Arab countries. Umer & Omer (2015) reviewed many researches that indicate teaching can be affected by testing either positively, negatively or even both, this depends on tests and tasks nature, poor test will results undesirable effect on both teaching and learning. All reviewed studies above agreed that tests and teaching cannot be separated, they strongly affect achieving the desired learning outcomes.

Before preparing any test educators must know:

- 1- Bloom Taxonomy to decide which learning area or domain (cognitive domain, affective domain, psychomotor domain) you want to examine.
- 2- Types of questions, their advantages and disadvantages.
- 3- Validity, reliability and their relation with leaning outcomes of the course.
- 4- Table of Specifications. (TOS)
- 5- Level of difficulties of the test questions.

A. Bloom Taxonomy

Bloom's Taxonomy was created by Benjamin Bloom 1965 in order to promote higher forms of thinking in education he developed it in 1979. He divided these Taxonomy to three learning domains:

- Cognitive Domain: test mental skills. (knowledge)
- Affective Domain: test growth in feelings or emotional areas. (attitude or self)
- Psychomotor Domain: test manual or physical skills. (skills)

Figure1 below illustrate the three learning domains.

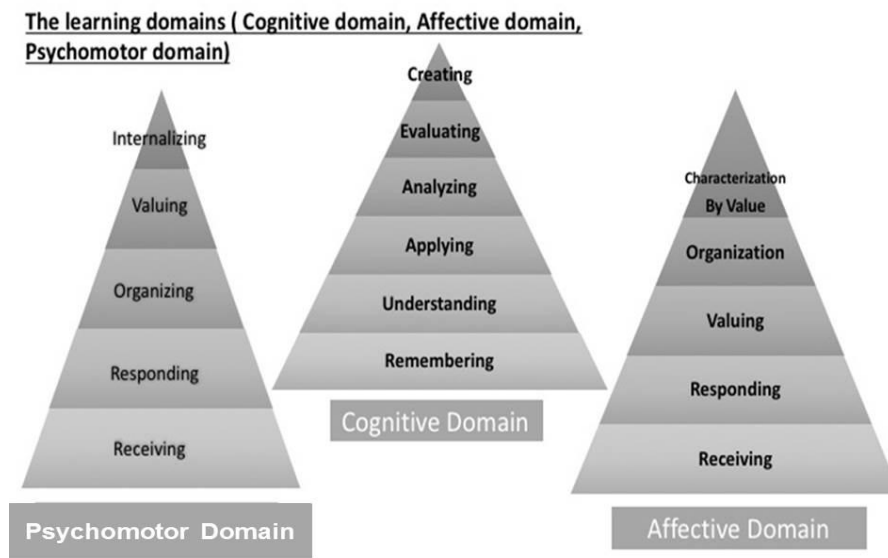


Figure.1 Bloom's Learning Domains
 Source: <https://www.papermasters.com/taxonomy-learning.html>

As cognitive domain is usually used to test language learning, this paper will focus on how to use the cognitive domain to make test questions. Benjamin Bloom and colleagues have suggested six levels of learning cognitive domain. These six levels are: remembering, understanding, applying, analyzing, evaluating and creating. To achieve the educational objectives Bloom classifies the educational process in the cognitive domain starting from the simplest behavior "remembering" to the most complex "creating". The first one must be mastered before the next one can take place to serve the development of intellectual abilities and skills.

Lorin Anderson and David Krathwohl; former students of Bloom; in 2000 rearranged the levels of the cognitive domain in the mid-nineties and made some changes in the higher levels , which perhaps the most noticeable ones being, see figure 2 below to recognize the change.

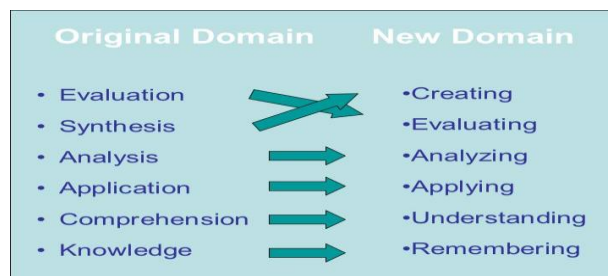


Figure.2 Original Taxonomy vs. Revised One
 Source: <http://www.nwlink.com/~donclark/hrd/bloom.html>

Each learning level in the cognitive domain is assigned for intended learning outcomes or objective .Figure 3 below illustrates the six cognitive domains learning levels, and their objective, it also illustrates the sequence and the development of knowledge and concepts through the educational process form the easiest level “remembering” to the most difficult level “creating”.

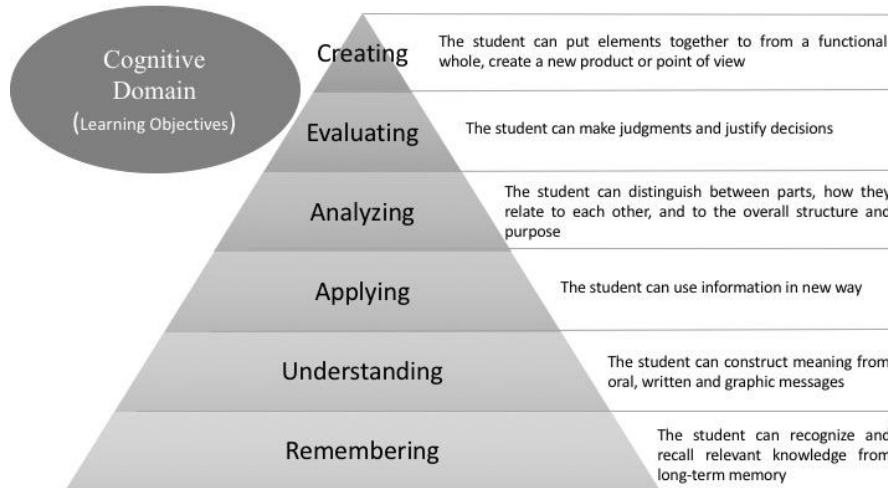


Figure 3: Cognitive Domain (Learning Objectives)
 Source: <http://edu4techs.com/archives/3069>

After learning, the objectives of each learning level of the cognitive domain, it is important to learn about stem verbs used to phrase learning objectives of each domain. The Learning and Teaching Unit at the University of New South Wales in Australia have presented a list of verbs to help educators to construct learning outcomes which align with Bloom Taxonomy. Figure 4 below illustrates some of the stem verbs that are used to write learning objectives of each level of the cognitive domain, and help teachers: 1. Map their teaching. 2. Achieve learning objectives of the cognitive domain. 3. Use the right stem verb in writing tests - questions.

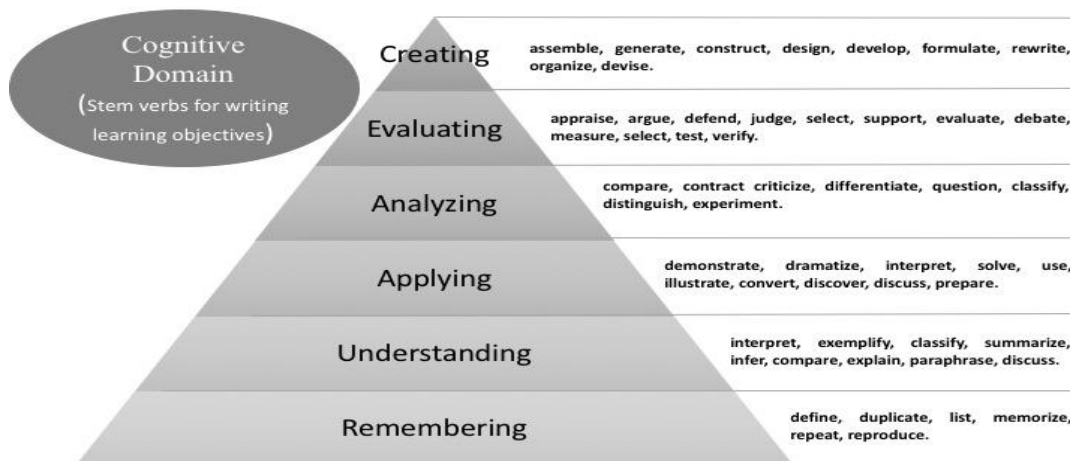


Figure 4: Cognitive Domain (Stem verbs for writing learning objectives)
 Source: <https://cft.vanderbilt.edu/guides-sub-pages/blooms-taxonomy/>

B. Types of Questions

Tracy (2015) pointed there are two types of tests questions, they can be objective or subjective questions, each type of these questions has its advantages and disadvantages which teachers should consider when they write their test; these advantages and disadvantages are not covered in this paper.

Writing test-questions on any given subject matter must be prepared according to the objectives and learning outcomes of the course, teachers must choose the best test format for the cognitive ability or comprehension that they seek to evaluate, and this format of test-questions must be suitable to achieve desired learning outcomes of course. So before writing a test, it is vital to specify; what do you want to test? what is the objective of the test? These questions will help you to choose the suitable type of questions; there are different types of test such as:

- A proficiency test is one that measures a candidate's overall ability in a language; it isn't related to a specific course.
- An achievement test on the other hand tests the students' knowledge of the material that has been taught on a course.
- A diagnostic test highlights the strong and weak points that a learner may have in a particular area.
- A prognostic test attempts to predict how a student will perform on a course.

Tests types mentioned above are constructed by objective or subjective questions; the table below illustrates main types of tests-questions.

TABLE I
TYPES OF TEST-QUESTIONS

Tests questions			
No.	A- Objective Questions	No.	B- Subjective Questions
1	Multiple choice questions. (MCQs)	1	Short answer questions
2	True/false questions	2	Essay questions: a- Guided writing questions. b- Open writing questions.
3	Matching questions	3	Performance test / Oral questions
4	Fill-in-the-Blank Questions		

C. *Validity, Reliability and Their Relation with Learning Outcomes of the Course*

To achieve learning outcomes of the course teachers should be aware about the following definitions," validity, reliability, goal, objective and learning outcome" these are connected lines in the educational process.

- Validity refers to whether or not the test measures what it claims to measure.
- Reliability has to do with the consistency, or reproducibility, or an examinee's performance on the test, for example, if you were to administer a test with high reliability to an examinee on two occasions, you would be very likely to reach the same conclusions about the examinee's performance both times.
- Goal – A goal is a broad definition of student competence that is planned to be achieved.
- Objective – A course objective describes what a faculty member will cover in a course to reach goals (teaching methods, materials, strategies).
- Student Learning Outcome – A detailed description of what a student must be able to do at the conclusion of a course, and it must be measurable criteria to help guiding teaching, learning, and assessment process of the course.

Each course specification has clear intended learning outcomes (*ILOs*) to be achieved at the end of the course; teachers should consider these (*ILOs*) when they write their tests' questions, results of these tests are the indicators of success or failure of achieving the course learning outcomes. Any teacher must consider the learning domain level that must be useful to assess the course objectives and learning outcomes, and then choose the suitable type of questions that cover the different parts of the course and this will be clear after reading table of specifications (TOS) section and understand the application rules for using it to construct questions.

D. *Table of Specifications. (TOS)*

Notar, Zuelke, Wilson & Yunker (2004) defined (TOS), which sometimes is called a test blueprint, is a table that helps teachers align objectives, instructions, and assessments. (TOS) is a two way chart which describes the topic to be covered by the test or exam and number of items which should be associated with each topic. Grondlund (2006) pointed that (TOS), sometimes called a test blueprint, is a table that helps teachers map a test onto their instructional objectives for a given segment of study. Wolming & Wilkstrom (2010) stated that the cornerstone of classroom assessment practices is the validity of the judgments about students' learning and knowledge and this validity can be achieved by using (TOS) which can be taken as a map that help teachers construct ideal test fulfilling course - objectives, not only must (TOS) be developed before the test is written, it must be developed before the actual teaching. Helenrose Fives & Nicole Di Donato-Barnes (2013) stated that classroom tests provide teachers with essential information used to make decisions about instruction and student grades. A table of specification (TOS) can be used to help teachers frame the decision making process of test construction and improve the validity of teachers' evaluations based on tests constructed for classroom use. In this article we explain the purpose of a (TOS) and how to use it to help construct classroom tests. Nicole, Helenrose and Emily S. (2013a) (2013b) investigated if instruction of Table of Specifications (TOS) would influence the quality of classroom test construction. They pointed that results proved (TOS) is informative for educational researchers, teacher educators, and practicing teachers interested in evidenced-based strategies that may improve assessment-related practices, and they found that participants were able to use (TOS) tool and choose items that reflect the subject matter specified in the (TOS) tool, however, they experienced difficulty selecting items at the cognitive level specified in the (TOS) tool. Helenrose & Nicole (2013a) (2013b) approved that (TOS) is one tool that teachers can use to support their professional judgment when creating or selecting test for use with their students. The (TOS) can be used in conjunction with lesson and unit planning to help teacher make clear the connections between planning, instruction, and assessment.

As a conclusion to the previous studies, table of specification (TOS) has many advantages; (TOS) can be used to help educators to:

- 1- Map a test into their instructional objectives for a given segment of study.
- 2- Make the decision about test making process and construction.
- 3- Improve the validity of teachers' evaluations based on tests constructed for classroom use.
- 4- Identifies the content area covered in the class.
- 5- Identifies the performance objectives at each level of the cognitive domain of Blooms Taxonomy.

Steps of using (TOS)

Abdelwahab , Al-Bashier , Abderahaman (2009) stated the steps of using (TOS) as follow:

- Specify the educational goals of the content which will be tested.
- Specify the items in the content to be assessed.
- Specify the percentages of the content division based on the following equation.

$$\frac{\text{Hours of each unit}}{\text{Total hours of the course}} \times 100$$

- Specify the level of the cognitive domain to be assessed according to the teaching specification. E.g. 25% remembering, 30% understanding, 20% applying, 20% creating.
- Specify the number of questions for the test.
- Specify the number of items in each question.

Table 2 below illustrate example for content analyze as a step to construct questions.

TABLE. 2
EXAMPLE OF HOW TO FIND THE PERCENTAGES OF THE CONTENT DIVISION ACCORDING TO COURSE SPECIFICATION BASE ON THE FOLLOWING EQUATION $\frac{\text{Hours of each unit}}{\text{Total hours of the course}} \times 100$

Content analysis according to course specification Grammar 2 ENG-3 (117)			
List of Topics	No. of Weeks	Contact hours	Percentages of the hours to the content to construct the test question
Connecting ideas	2	6	$\frac{6}{45} \times 100 = 13\%$
Comparisons	2	6	$\frac{6}{45} \times 100 = 13\%$
The passive	2	6	$\frac{6}{45} \times 100 = 13\%$
Count/Non-count nouns and articles	2	6	$\frac{6}{45} \times 100 = 13\%$
Adjective Clauses	2	6	$\frac{6}{45} \times 100 = 13\%$
Gerund and infinitives	2	6	$\frac{6}{45} \times 100 = 13\%$
Noun Clauses	2	6	$\frac{6}{45} \times 100 = 13\%$
Revision	1	3	$\frac{3}{45} \times 100 = 7\%$
Total hours	15	45	

After specifying the percentages of the content division; which indicate the percentage of each part of the content to the total percentage of the test-questions; then teachers move to the next step to find which level of cognitive domain to be assessed and number of questions for each level in that part according to the division results this can be illustrated in table 3.

TABLE 3
EXAMPLE OF USING (TOS) TO CREATE QUESTIONS SHOWING THE NUMBER OF EACH PART OF THE CONTENT DIVIDED TO THE LEARNING LEVEL ACCORDING TO THE COURSE OBJECTIVES

Example of using TOS to create questions							
Content	Percentage in the test	Remembering	Understanding	Applying	Evaluating	Creating	Total questions
Connecting ideas	20%	3		2	2	3	$\frac{20}{100} \times 50 = 10$ q
Comparisons	20%		3	2	2	3	$\frac{20}{100} \times 50 = 10$ q
The passive	20%	2	3		2	3	$\frac{20}{100} \times 50 = 10$ q
Count/Non-count nouns and articles	20%	2	3	2	3		$\frac{20}{100} \times 50 = 10$ q
Adjective Clauses	20%			5	5		$\frac{20}{100} \times 50 = 10$ q
Total question of the test	100						50

The above table illustrates how to divide questions of the content according to the time devoted for teaching each part of the course unit, and also helps you draw a map of which level of learning domain you need to test. If we take connecting idea unit as example, there will be 10 questions out of 50 question in the test in this area, then according to the objectives, 3 questions test remembering, 2 questions test applying, 2 questions that test evaluating, 3 questions that test creating, to be total 10 questions that test connecting idea unit. If we take "The Passive" unit, there should be 10 questions, 3 questions test creating, 2 questions test evaluating, 3 questions test understanding and 2 questions remembering, and so on for the rest of the units using same equation.

D. Level of Difficulties of the Test Questions

Level of difficulties of the test questions can be known by applying the test questions in a small number of students as a sample, then analyzing the results according to the marks.

There are many advantages of analyzing exam questions:

- 1- Gives the teacher feedback about questions to be developed.
- 2- Draw a map-picture about planed skills students didn't achieve.
- 3- Help teachers to evaluate themselves.
- 4- Build the foundation to create questions bank.
- 5- Help to develop better tests that have certain characteristic of validity and reliability and objectives in the future.

How to evaluate level of difficulties for test- questions?

We can do that basing on the following equation to find the percentage that reflects level of difficulties.

$$\text{Indictor of difficulties} = \frac{\text{no.of correct answers}}{\text{total no.of students}} \times 100$$

Table 4 below illustrates percentage agreement to questions' level of difficulties.

TABLE IV
PERCENTAGE AGREEMENT TO QUESTIONS' LEVEL OF DIFFICULTIES.

Difficult	Ideal	Easy
Less than 50%	50%	More than 50%

If we take the percentage in the table above in example: E.g. If 30 students sat for the test, 10 students only answer the 1st question then we can calculate this test difficulties of that question based on the equation mentioned above as follow:

$$\text{Level of difficulties} = \frac{10}{30} \times 100 = 33\% \text{ and that means it is difficult}$$

Table .5 below illustrates questions' level of difficulties according to percentages after getting the results doing the equation $\frac{\text{no.of correct answers}}{\text{total no.of students}} \times 100$

TABLE. V
QUESTIONS' LEVEL OF DIFFICULTIES ACCORDING TO PERCENTAGES

Question No.	Indicator of level of difficulties	meaning
1	25%	Difficult
2	70%	Easy
3	60%	Easy
4	90%	Very easy
5	50 %	Ideal
6	5%	Very difficult

Classifying questions-level of difficulties help to reuse the questions in future test if it is stored in a designed questions- bank.

III. METHODOLOGY

The descriptive analytical method is used to conduct this study.

A. Data Collection

The study has been carried in King Khalid University, English Department, College of Science & Arts Muhayil in the academic year 2018-2019

B. Participants

English Language instructors of College of Science & Arts Muhayil are the participants of this research.

C. Research Instrument

- 10 testes are analyzed.
- A questionnaire is sent to English Language instructors in College of Science & Arts Muhayil 22 instructors out of 56 answered the questionnaire.

D. Validity and Reliability of the Questionnaire

According to table using Cronbach's Alpha, Spearman&Brown and - One-Sample Test it is obvious that .873 and .859 reflect the validity and reliability of the questionnaire is suitable for the study

TABLE. 6
VALIDITY AND RELIABILITY OF THE QUESTIONNAIRE

	Cronbach's Alpha	Spearman&Brown
Reliable	0.763	0.739
Validity= $\sqrt{\text{Reliable}}$	0.873	0.859

IV. DATA ANALYSIS AND DISCUSSION

A. *Tests*

The researcher have analyzed 10 test papers for English language courses, these tests were chosen randomly, two for Grammar, two for Applied Linguistics, two for Vocabulary Building, one for Speech Workshop, one for Reading, one for Research Methods and one for Study Skills. The researcher found that only three of these tests have covered 90% of the units that are part of the course specification, and only five of them consist of questions test higher levels of thinking of Bloom Taxonomy. This means instructors either do not use Bloom Taxonomy Cognitive Domain and Table of Specification to design and write their test questions, or they misuse them, this agrees with Nicole, Helenrose and Emily S. (2013) and Alexis & Ricardo (2009) studies which indicate that instructors need training to use Bloom Taxonomy and Table of Specification correctly in test writing as most researches agreed that testing and teaching are integrated and can't be teased .

B. *Questionnaire*

Twenty two instructors from English department answered the questionnaire.

1. General information

A. Gender: seven males instructors and fifteen are female instructors are the participant of this research paper.

B. Qualification: Seven instructors are Ph. D holder and 15 are Master Degree holder.

C. Experience: Eleven instructors have been teaching more than ten years, eight instructors have been teaching between 6-10 years and only, only three instructors have less than five years-experience. These results show that most of the participants have quite long experience of teaching.

2. Questionnaire questions analysis.

Table .6 illustrates questionnaire statistical results using SPSS

TABLE VII
QUESTIONNAIRE ANALYSIS.

statement	Mean	Std. Deviation	Test Value = 2			reality Of 0.05 significance	Value
			T	df	Sig		
1-Do you have any experience about Bloom Taxonomy domains?	1.59	.796	-2.409-	21	.025	significance	no
2- Do you plan your lesson using Bloom Taxonomy domains?	2.18	.907	.940	21	.358	insignificance	No idea
3- Do you use Bloom Taxonomy domains when you write your test questions?	2.32	.839	1.779	21	.090	insignificance	yes
4- Do you use the Affective domain of Bloom to write English test?	1.18	.588	-6.521-	21	.001	significance	no
5- Do you use the Cognitive Domain of Bloom to write English test?	2.09	1.019	.418	21	.680	insignificance	No idea
6- Do you use the Psychomotor Domain of Bloom to write English test?	1.73	.985	-1.299-	21	.208	insignificance	No idea
8- Did you find Bloom Taxonomy domains helpful to write test questions?	2.64	.492	6.062	21	.001	significance	yes
9- Did you get any training to use Bloom Taxonomy domains to write your test questions?	1.73	.935	-1.368-	21	.186	insignificance	No idea
10- Do you have any experience of using Table of Specifications?	2.09	.971	.439	21	.665	insignificance	No idea
11-Do you plan your course your course using Table of Specification?	1.86	.941	-.680-	21	.504	insignificance	No idea
12- Do you use Table of Specification to write your tests questions?	1.95	.999	-.213-	21	.833	insignificance	No idea
13-Did you get any training to use Table of Specifications to write your test?	1.45	.858	-2.982-	21	.007	significance	no
14-Did you find Bloom Taxonomy and Table of Specifications helpful to write test questions?	1.73	.883	-1.449-	21	.162	insignificance	No idea
15-Do you need training to use Bloom Taxonomy and Table of Specifications to write test questions?	1.86	.990	-.646-	21	.525	insignificance	No idea

1. Showing results of phrase No .1 “Do you have any experience about Bloom Taxonomy domains?” Seen from the table above, (T.test) value calculated is (-2.409-), with degree of freedom (21), and with a probability value (.025), which means that there statistical significance of the statistical reality conclusion of the term that says: the participants have no experience about bloom Taxonomy domain, it is clear from the reality of statistical inference and the approval subjects of this phrase moderately at the significance level of (0.05)
2. Showing results of phrase No .2 “Do you plan your lesson using Bloom Taxonomy domains?” Seen from the table above,(T.test) value calculated is (.940), with degree of freedom (21) ,and with a probability value (.358), which means that the statement is not significant, and the statistical reality conclusion of the term that says: the participant don’t plan their lesson using Bloom Taxonomy domains ,this is clear from the reality of statistical inference which is not approved by significance level of (0.05)
3. Showing results of phrase No .3 “Do you use Bloom Taxonomy domains when you write your test questions?” Seen from the table above, (T.test) value calculated is (1.779), with degree of freedom (21), and with a probability value (.090), the statement is not significant, and the statistical reality conclusion of the term that says: the participant do not use Bloom Taxonomy domains to write their test – questions, it is clear from the reality of statistical inference approval subjects of this phrase moderately at the significance level of (0.05)
4. Showing results of phrase No (4) “Do you use the Affective domain of Bloom to write English test?” “Seen from the table above, (T.test) value calculated is (-6.521-), with degree of freedom (21), and with a probability value (.001), which means that there statistical significance of the statistical reality conclusion of the term, that says: the participant don’t use the Affective domain of Bloom to write English test. This is clear from the reality of statistical inference approval subjects of this phrase moderately at the significance level of (0.05)

5. Showing results of phrase No (5) “Do you use the cognitive domain of Bloom to write English test?” Seen from the table above, (T.test) value calculated is (.418) , with degree of freedom (21) ,and with a probability value (.680) , which means that the statement is not significant, and the statistical reality conclusion of the term, shows that: the participant are not sure if they use the cognitive domain of Bloom or not to write English test . It is clear from the reality of statistical inference which is not approved by significance level of (0.05)
6. Showing results of phrase No (6) “Do you use the Psychomotor domain of Bloom to write English test?” Seen from the table above, (T.test) value calculated is (-1.299-), with degree of freedom (21), and with a probability value (.208), which means that the statement is not significant, and the statistical reality conclusion of the term says that: the participant are not sure if they use the Psychomotor domain of Bloom or not to write English test. This is clear from the reality of statistical inference which is not approved by significance level of (0.05)
7. Showing results of phrase No (6) *Which level of Bloom Taxonomy do you use to write English tests? Tick all levels you use.* Results can be read from the figure 6 below, which illustrates that no one use creating level in writing question, and many ticked combination of levels, but not more than three levels that are used together, twelve participants use remembering level, twelve participants use understanding level, thirteen participants use applying level, thirteen participants use analyzing level, and six participants use evaluating level and no one use creating in their tests.

Figure .6 results reflects the frustration that instructors have about using levels of Bloom Taxonomy in their test, which means either they use it wrongly or don't have experience about using it .

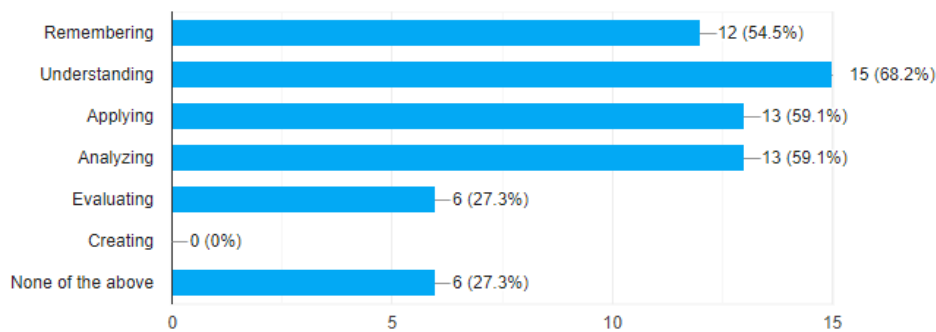


Figure 5. Participants' answers to question .7

8. Showing results of phrase No (8) “Did you find Bloom Taxonomy domains helpful to write test questions?” Seen from the table .6 , (T.test) value calculated is (6.062) , with degree of freedom (21) ,and with a probability value (.001), which means that there is a statistical significance of the statistical reality conclusion of the term that says: the participant agreed that Bloom Taxonomy domains are helpful to write test questions. It is clear from the reality of statistical inference approval subjects of this phrase moderately at the significance level of (0.05)
9. Showing results of phrase No (9) “Did you get any training to use Bloom Taxonomy domains to write your test questions?” Seen from the table above, (T.test) value calculated is (-1.368-), with degree of freedom (21), and with a probability value (.186) , which means that the statement is not significant , and the statistical reality conclusion of the term that says: the participant did not get any training to use Bloom Taxonomy domains to write your test questions. This is clear from the reality of statistical inference which is not approved by significance level of (0.05)
10. Showing results of phrase No (10) “Do you have any experience of using Table of Specifications?” “Seen from the table above, (T.test) value calculated is (.439-), with degree of freedom (21), and with a probability value (.665), which means that the statement is not significant, and the statistical reality conclusion of the term that says: the participant do not have any experience of using Table of Specifications. It is clear from the reality of statistical inference which is not approved by significance level of (0.05)
11. Showing results of phrase No (11) “Do you plan your course your course using Table of Specification?” Seen from the table .6, (T.test) value calculated is (-.680-), with degree of freedom (21), and with a probability value (.504), which means that the statement is not significant, and the statistical reality conclusion of the term that says: the participant do not have clear answer if they plan their course using Table of Specification .This is clear from the reality of statistical inference which is not approved by significance level of (0.05)
12. Showing results of phrase No (12) “Do you use Table of Specification to write your tests questions?” Seen from the table .6, (T.test) value calculated is (-.213-) ,with degree of freedom (21), and with a probability value (.833) , which means that the statement is not significant, and the statistical reality conclusion of the term that says: the participant do not have clear answer about using Table of Specification in writing their test using . It is clear from the reality of statistical inference which is not approved by significance level of (0.05)

13. Showing results of phrase No (13) which reads “Did you get any training to use Table of Specifications to write your test?” Seen from the table above, (T.test) value calculated is (-2.982-), with degree of freedom (21), and with a probability value (.007), which means that there statistical significance of the statistical reality conclusion of the term that says: the participant did not get any training to use Table of Specifications to write their tests. This is clear from the reality of statistical inference approval subjects of this phrase moderately at the significance level of (0.05)
14. Showing results of phrase No (14) “Did you find Bloom Taxonomy and Table of Specifications helpful to write test questions?” Seen from the table .6, the (T.test) value calculated is (-1.449-), with degree of freedom (21), and with a probability value (.162), which means that the statement is not significant, and the statistical reality conclusion of the term that says: the participant were not sure if Bloom Taxonomy and Table of Specifications are helpful in writing test questions .This is clear from the reality of statistical inference which is not approved by significance level of (0.05)
15. Showing results of phrase No (15) “-Do you need training to use Bloom Taxonomy and Table of Specifications to write test questions?” Seen from the table .6, (T.test) value calculated is (-.646-), with degree of freedom (21), and with a probability value (.525), which means that the statement is not significant, and the statistical reality conclusion of the term that says: the participant don’t know if they need training or not to use Bloom Taxonomy and Table of Specifications to write test questions. This is clear from the reality of statistical inference which is not approved by significance level of (0.05)
16. Showing results of phrase No (16) Do you think Bloom Taxonomy and Table of Specifications are important to write test questions? Explain. Eleven of the participant answered positively while seven of them answered negatively and four participants their answers were neutral.

V. RESULTS AND DISCUSSION

We can see that most of the participants have more than six years’ experience in teaching regardless to the gender , and most of them are master degree holders , question number one results show that the participants have no experience about Bloom Taxonomy domain .Participants’ answer of question two and three prove that participant don’t plan their lesson or write their tests question using Bloom Taxonomy domains ,which reflect a dangerous feature in the educational field .Participants’ answers of question four show that they don’t use the Affective domain of Bloom Taxonomy to write English test , also results of question five and six say that the participants are not sure if they use the cognitive and Psychomotor domains of Bloom , this reflects participants’ ignorance about Bloom Taxonomy in the educational field. Participants’ answers of question seven about levels of Bloom Taxonomy which they use in their test reflect participants’ confusion, that even if they use Bloom Taxonomy in writing question, they use it wrongly. Answers of question eight showed that participants agree that Bloom Taxonomy domains are helpful to write test questions in spite of the confusion that they have. Participants’ answers of question nine reflect the reason behind participants’ confusion about Bloom Taxonomy domains, is that they didn’t get training to use Bloom Taxonomy domains in their lesson planning and test-questions writing. Participants’ answers of question ten show that participants do not have any experience of using Table of Specifications, which is a necessity to test- questions balance. Participants’ answers of questions eleven and twelve prove that participants do not have clear answer if they plan, write their questions using Table of Specification or not. Participants’ answers of question thirteen show that they didn’t get training to use Table of Specification questions. Answers of Question fourteen show that participants were not sure if Bloom Taxonomy and Table of Specifications are helpful to write test questions. Participants’ answers of question fifteen show that they cannot recognize their need for a training to use Bloom Taxonomy and Table of Specifications in writing test questions. Participants’ answers of question sixteen reflect a clear picture of the confusion the participants have about Bloom Taxonomy and Table of Specifications, as some said they have heard about them, but didn’t know their importance and how to use them ,while others explain that they have learnt about them and used them in lessons-planning and test question writing , although this question answers doesn’t agree with participants answers of question seven, which show that most of the participants don’t use more than three levels of Bloom Taxonomy in their tests-questions.

Although a few participants’ opinions reflect their knowledge of both Bloom Taxonomy and Table of Specifications, still the results show that participants need training to use Bloom Taxonomy and Table of Specification in their lesson planning and test-writing.

Results of question eight agree with McMillan (2003) study that assessment is still based on traditional behaviorist theories of learning.

Results of question seven, eleven and thirteen agree with Alexis, & Ricardo (2009) study that teachers have a negative view of assessment, and need adequate training in language assessment for all prospective language teachers is important.

Results of questions twelve, thirteen and fourteen which show that participant need training to use Table of Specification are clarified by Helenrose & Nicole (2013a) (2013b) who said TOS is a connections between planning, instruction, and assessment. Results must be considered, taking Wolming & Wilkstrom (2010) result , (TOS) which can be taken as a map that help teachers construct ideal test fulfilling course.

VI. CONCLUSION

Regarding the results above, reasons behind questions' weakness of English Language tests are the inadequate way of designing tests, as a result of ignorance of basics of test questions criteria and lacking main knowledge of how to use Bloom Taxonomy and Table of Specification.

Instructors need a training to use Bloom Taxonomy and Table of Specification, to have the basics of designing Tests' questions of English Language, that reflects (ILOs) in a pedagogical way. Learning these basis will help them develop their test in scientific pedagogical way.

Training will help instructors:

- Develop good questions that suitable to all students regardless to students' different personality or learning style.
- Evaluate and develop their questions standard by adapting it according to the results of the test .
- Prepare ideal test questions according to desired objectives of the unit and the course.
- Keep teaching and testing integrated and interwoven in a scientific way.
- Help teachers feel secure to have their questions ready in any situation.
- Keep teaching and testing integrated and interwoven in a scientific way
- Decrease the stress accompanied with test time for teachers.

APPENDIX. INSTRUCTORS QUESTIONNAIRE

Dear colleague,

This questionnaire is an instrument to collect data of a research paper titled:

“Basis of Designing Test Questions of English Language That Reflect Intended Learning Outcomes (ILOs)”

The researcher is trying to find out reasons behind questions' weakness of English language tests and solutions that can help teachers use a scientific basis for writing ideal tests-questions in a pedagogical way that tests all course content, and reflects students' learning outcomes. Your answers will be confidential and will be used only for research purpose.

Thank you for your cooperation.

Dr. Mona M. Hamad

1-General information

- A. Gender** : Female Male
B. Qualification: Bachelor Master Doctorate
C. Your experience in teaching field: 0-5 years 6-10 years more than 10 years

2-Questionnaire

	No	Question	Yes	No Idea	No
1		Do you have any experience about Bloom Taxonomy domains?			
2		Do you plan your lesson using Bloom Taxonomy domains?			
3		Do you use Bloom Taxonomy domains when you write your test questions?			
4		Do you use the Affective domain of Bloom to write English test?			
5		Do you use the Cognitive Domain of Bloom to write English test?			
6		Do you use the Psychomotor Domain of Bloom to write English test?			
7		Which level of Bloom Taxonomy do you use to write English tests? Tick all levels you use.			
		• Remembering			
		• Understanding			
		• Applying			
		• Analyzing			
		• Creating			
		• Evaluating			
		• None of the above			
8		Did you find Bloom Taxonomy domains helpful to write test questions?			
9		Did you get any training to use Bloom Taxonomy domains to write your test questions?			
10		Do you have any experience of using Table of Specifications			
11		Do you plan your course your Table of Specification?			
12		Do you use Table of Specification to write your tests questions?			
13		Did you get any training to use Table of Specifications to write your test?			
14		Did you find Bloom Taxonomy and Table of Specifications helpful to			

- write test questions?
- 15 Do you need training to use Bloom Taxonomy and Table of Specifications to write test questions?
- 16 Do you think Bloom Taxonomy and Table of Specifications are important to write test questions? Explain.

REFERENCES

- [1] Abdelwahab, A., Al-Bashier, M., Abderahaman, N. (2009). Guide-book of Achievement Test Preparation, Khartoum p 52-59
- [2] Alexis, A., Lopez, M., & Ricardo, B. (2009). Language Testing in Colombia: A Call for More Teacher Education and Teacher Training in Language Assessment, profile Vol. 11, No. 2, p.55, 70.
- [3] Bloom, B. (1965). Taxonomy of Educational Objectives, Handbook, The Cognitive Domain, New York : David McKay Co. Inc. P.18
- [4] Bloom's Taxonomy of Learning Domains .Retrieved from <http://www.nwlink.com/~donclark/hrd/bloom.html> accessed 12/2/2019.
- [5] Bloom's Taxonomy of Educational Objectives and Writing Intended Learning Outcomes Statements Retrieved on from: https://www.stonybrook.edu/commcms/provost/assessment/_documents/blooms-taxonomy_IACBE.pdf (accessed 12/2/2019).
- [6] Brown, H. (1994). teaching by Principles an Interactive Approach to Language Pedagogy , Prentice-Hall regent, New Jersey p. 375.
- [7] Brown, H. (2004). Language assessment: Principles and classroom practices, London: Longman p.6
- [8] Grondlund, N. (2006). Assessment of student achievement (8th ed.). Boston, MA: Pearson.
- [9] Heather, F., Steve, K., Stephanie, M. (2009) A Handbook for Teaching and Learning in Higher Education Enhancing Academic Practice , third edition, Routledge , Taylor and Francis group,p.44-46.
- [10] Helenrose, F.& Nicole, D. (2013). Classroom Test Construction: The Power of a Table of Specifications, Practical Assessment, Research & Evaluation journal ,Volume 18, Number 3. Retrieved from: <http://pareonline.net/getvn.asp?v=18%26n=3> (accessed 12/2/2019).
- [11] How to write Good Test Questions, Help Teaching, Retrieved from: (accessed 12/2/2019)https://www.help-teaching.com/about/how_to_write_good_test_questions.
- [12] How do you Determine if a Test has Validity, Reliability, Fairness, and Legal Defensibility? Professional Testing Inc. © PTI 2006 Retrieved from:http://www.proftesting.com/test_topics/pdfs/test_quality.pdf (accessed 12/2/2019).
- [13] John, B. & Catherine, T. (2007). Teaching for Quality Learning at University , What the Student Does , 3rd edition McGraw-Hill Companies p. 195- 197.
- [14] Learning Outcomes, Essay as ABCD, CBU, retrieved from https://www.cbu.edu/assets/2091/writing_perfect_learning_outcomes.pdf (accessed 12/2/2019).
- [15] Learning outcomes: specifying levels of academic performance, Retrieved from:https://teaching.unsw.edu.au/sites/default/files/upload-files/outcomes_levels.pdf(accessed 12/2/2019)
- [16] McMillan, J. (2003). Understanding and improving teachers' Classroom Assessment Decision Making: Implications for theory and practice. Educational Measurement: Issues and Practice, 22(4), 34-43.
- [17] Notar, C., Zuelke, D., Wilson, J., Yunker, B. (2004). The Table of specifications: Insuring accountability in Teacher Made Test, Journal of Instructional Psychology Jun vol. 31 No .2 p. 115-129.
- [18] Nicole D., Barnes, H., Fives & Emily, S. (2013) ,Using a Table of Specifications to improve teacher-constructed traditional tests: an experimental design, Assessment in Education: Principles, Policy & Practice, Vol. 21, No. 1, 90–108, Retrieved from : <http://dx.doi.org/10.1080/0969594X.2013.808173> (accessed 12/2/2019).
- [19] Questioning STEM bookmark, Stem learning .Retrieved from <https://www.stem.org.uk/resources/elibrary/resource/273282/questioning-stem-bookmark> (accessed 12/2/2019).
- [20] Reynolds, C., Livingston, R., & Willson, V. (2006). Measurement and assessment in education. Boston, MA: Allyn and Bacon.
- [21] Test Questions Types, Teaching English , British Council, Retrieved from <https://www.teachingenglish.org.uk/article/test-question-types> (accessed 12/2/2019)
- [22] Tracy, J. (2015). Objective or Subjective? Those are the Questions, Palomar college , learning for success ,Retrieved from: <https://www2.palomar.edu/pages/testwritingstrategies/2015/04/13/objective-or-subjective-those-are-the-questions>(accessed 12/2/2019).
- [23] Umer, M. & Omer, A. (2015). An Investigation of Saudi English-Major Learners' Perceptions of Formative Assessment Tasks and Their Learning, English Language Teaching; Vol. 8, No. 2; 2015 Retrieved from URL: <http://dx.doi.org/10.5539/elt.v8n2p109> (accessed 12/2/2019).
- [24] Wiggins, G. (1992). Creating tests worth taking. *Educational Leadership*, 26, 26-33.
- [25] Wolming, S. & Wikstrom, C. (2010). The concept of validity in theory and practice. *Assessment in Education: Principles, Policy & Practice*, 17, 117-132.

Mona M. Hamad is a Sudanese, her Ph.D. in Education “Curriculum & Teaching Methods of English Language” from Al-ZaiemAzhhri University, Khartoum, Sudan. She obtained her Master degree in Education Technology (2005) from Al-ZaiemAzhhri University, Khartoum, Sudan. She also obtained a higher degree Diploma in ELT (2001) from Khartoum University, Khartoum, Sudan. Her Bachelor of Arts (1991) was obtained from AL-Neelain University, Khartoum, Sudan.

She has been working as Assistant Professor at King Khalid University, Muhayil, English Department since 2012, She has worked as Part-time Assistant Professor for Al-ZaiemAzhri University , Sudan, Khartoum from 2010-2012 , she also worked as English language teacher for Secondary level for the Ministry of General Education , Sudan from 1996-2012 .

She published many papers and two books titled:

1. Revision papers for Secondary Students, Khartoum, Sudan
2. Special Teaching Methods, Khartoum, Al-ZaiemAzhri University, Sudan.

Dr. Mona M. Hamad is a member of KSAALT /TESOL (Kingdom of Saudi Arabia Association of English Language Teachers) & ASTEL (Association of Sudanese Teachers of English Language).