



Journal of Tertiary Education and Learning (JTEL)

ISSN: 2994-4015 (ONLINE)

VOLUME 3 ISSUE 2 (2025)



PUBLISHED BY
E-PALLI PUBLISHERS, DELAWARE, USA

Evaluating Iraqi EFL University Teacher-Made Tests in Light of the Courses' Objectives

Tabarak Mahmoud Ali¹, Sarah Mahmoud Ali^{2*}, Amani Salih Znad¹

Article Information

Received: July 15, 2024

Accepted: August 21, 2024

Published: April 21, 2025

Keywords

Courses' Objectives, EFL University Teacher-Made Tests

ABSTRACT

Testing in education is considered a crucial factor and a systematic tool in assessing students' success or failure. A good test must consist of various levels in line with Blooms' taxonomy in order to reflect the different capabilities of students. The current study aims at evaluating Iraqi EFL university teacher-made tests in light of courses' objectives. To achieve the purpose of the study, four Iraqi EFL university teacher-made tests of the final exams in the department of English at the College of Education for Women have been chosen to be the sample of the present study. For collecting data, the general objectives of the included courses have been identified; the content of the courses has been analyzed; the behavioral objectives have been derived; the list of the action verbs have been adopted and applied; then, the teacher-made tests have been analyzed according to the test analysis format. The findings indicate that Iraqi EFL university teacher-made tests at the College of Education for Women do not reflect the courses' objectives.

INTRODUCTION

Evaluation in English language teaching is a process of collecting, analyzing and interpreting information about teaching and learning in order to make informed decisions that enhance success of educational programs and students' achievement (Genesee & Upshur, 1996).

According to Boud and Falchikov (2007), assessment has a powerful effect on what students do and how they do it. It informs them what they can and cannot succeed in doing. Moreover, it builds or undermines their confidence - as learners on a course and in the future, and in the world.

There are two major categories of tests: standardized and non-standardized. Standardized tests are administrated and scored using fairly uniform standards. Non-standardized or teacher-made tests are tests that constructed by the classroom teachers (Wiley, 2008). Teacher-made tests are probably the most commonly used way to assess students' achievement and progress in specific content area (Fagan & Spurrell, 1995). These tests, like all other types of assessments must be reliable and valid. The construction of a teacher-made test must be related directly to the course objectives and classroom instructions (Nath & Cohen, 2011).

Gorsuch and Griffee (2017) define the course objectives as "statements of how the course purpose can be accomplished". They state that test items should reflect the teaching or instructional objectives of each course. Otherwise, course objectives run the risk of being stated in an abstract and unverifiable way. Therefore, teachers stipulate how they would know the extent to which that objective had been accomplished by connecting a test or test items to a particular course objective.

Moreover, assessment plays an indispensable role in the educational process. It identifies whether the teacher has achieved the objectives he/she sets or not. The behavioral objectives provide the teachers with the idea

of what is the best assessment tool to be used in a given situation. Furthermore, they determine the behavior to be measured. Therefore, without instructional objectives, the teachers will have no indicator to measure achievement (Magazine, 2018). Students can, with difficulty, escape from the effects of poor teaching, but they cannot escape from the effects of poor assessment (Boud, 1995).

The problem of the study is best expressed through answering the following questions:

1. Are Iraqi EFL university teacher-made tests at the College of Education for Women designed and developed according to the courses' objectives?
2. How adequate are Iraqi EFL university teacher-made tests items in reflecting the six cognitive levels of objectives?
3. To what extent do Iraqi EFL university teacher-made tests items measure thinking at lower and higher levels?
4. How adequate are Iraqi EFL university teacher-made tests employ the action verbs in their test items?

Aim

This study aims at evaluating Iraqi EFL university teacher-made tests in light of courses' objectives.

Limits

The study is limited to Teacher-made tests of the final exams/ first attempt at the English department of the University of Baghdad (College of Education for Women) during the academic year 2022-2023. It is also restricted to the revised version of Bloom's taxonomy -the cognitive domain, 2001.

Value

It is hoped that the present study will help in improving educational practices and programs by shedding light

¹ University of Information Technology and Communications, Baghdad, Iraq

² The Iraqi University, Baghdad, Iraq

* Corresponding author's e-mail: lampard4ever8@yahoo.com

on the importance of the courses' objectives; providing EFL university teachers with adequate knowledge and information regarding what it takes to have a good EFL test that will measure the six levels of the cognitive abilities; and in supplying curriculum developers and test designers with the solutions to the problem of testing by probably training teachers on how to construct classroom tests.

LITERATURE REVIEW

Language Testing

Testing is a method of measuring a learner's ability or knowledge or performance in a given domain (Brown, 2001). It is an important part of every teaching and learning experience (Madsen, 1988). Also, it is a curtail and fundamental part of the teaching-learning process (Ramos & Castillo, 2024). Like all educational assessment, language testing is a complex social phenomenon. It has evolved to fulfill a number of functions in the classroom, and society at large (Fulcher, 2010).

Heaton, (1999) mentions that tests provide teachers with useful feedback about what students have and have not learned. Knowledge of what students know and do not know is very important to the teachers and a lack of this knowledge is likely to bring about ineffective instruction. If teachers can make good use of test information, better teaching can be achieved. Feedback from tests can help the teacher provides more appropriate instructional guidance for individual students as well as for the class as a whole.

Teacher-Made Tests

Teacher-made tests (henceforth TMTs) are written or oral assessments that are not commercially produced or standardized (Board of Education for the City of Etobicoke, 1987). In other words, they are usually built without careful and thorough analyses and considerations of how the test should be (Tuckman, 1975). TMTs are designed by the teacher to meet his or her own course objectives. The teacher plans the test items to fit the instructional objectives.

The Planning Of Teacher-Made Tests

The importance of planning the TMT cannot be neglected. Patel (2011) suggests the following steps as a brief guide to the planning of tests of English as a forging language, especially when the teacher attends to measure students' achievement.

- Step 1:** Determine the Purpose of the Test
- Step 2:** What Teacher Wants to Measure
- Step 3:** Specifying the Course Content
- Step 4:** Specifying the Major Course Objectives
- Step 5:** Defining Behavioral Objectives
- Step 6:** Preparing a Table of Specification
- Step 7:** Selecting the Appropriate Item Format

Objectives

An objective is a point or an end view of the possible achievement in terms of what a student is able to do when the educational system is directed towards educational aims (Patel, 2011). Gronlund, (1999) on the other hand, mentions that teachers use the learning objectives for different purposes such as; providing a focus for instruction, providing guidelines for learning and providing targets for assessment.

The Revised Version of Bloom's Taxonomy (Anderson & Krathwohl Taxonomy, 2001)

During the 1990s, a new group of cognitive psychologists, led by Anderson, have updated the original version of Bloom's taxonomy (1956) to reflect relevance to twenty-first-century work (Bartlett, 2015). Anderson and his group have worked for five years to clarify the taxonomy (Stobaugh, 2013). The Revised Taxonomy has two dimensions: the knowledge dimension and the cognitive process dimension.

The structure of the knowledge dimension includes the factual knowledge, the conceptual knowledge, procedural knowledge and Meta-cognitive knowledge. Whereas, the cognitive process dimension of the new version of Bloom's taxonomy contains six levels which are based on the hieratical form and which move from the lowest level of thinking to the highest level of cognition (Bartlett, 2015). They are remembering, understanding, applying, analyzing, evaluating and creating

MATERIALS AND METHOD

Relying on the nature of the current study, the methodology involved is a qualitative. A qualitative methodology is defined by Gay *et al* (2012) as "the collection, analysis and interpretation of comprehensive narrative and visual data to gain insight into a particular phenomenon of interest."

Population and Sample

The population of the current study consists of Iraqi EFL University TMTs in the department of English, College of Education for Women of the University of Baghdad. While the sample of the study is Iraqi EFL University TMTs of the final exams, the first attempt during the academic year 2022-2023. The TMTs have been chosen from the four stages and one course from each stage. Thus, the TMTs for the following courses which have been included in the sample have been selected randomly:

1. First stage: Phonetics and Phonology;
2. Second stage: Short Story;
3. Third stage: Writing Academic Essay; and
4. Fourth stage: Drama

The total number of teacher-made tests' sample is (4) and they include (122 items). The distribution of the tests' items is shown in Table 1.

Table 1: The Distribution of the Tests' Items

Stage	Courses	Number of Items
First	Phonetics and Phonology	83
Second	Short story	7
Third	Writing Academic Essay	24
Fourth	Drama	8
Total		122

The Instrument of the Study

In order to collect data, content analysis have been used as a research instrument. Wilkinson and Birmingham (2003) believe that content analysis is a powerful research instrument to decide, from the content of a message or a text, sensible findings concerning the work of the writer or an author. It could be used successfully to analyze texts and solve issues. Thus, to evaluate TMTs in light of courses' objectives, a test's analysis format has been constructed in light of the course objectives. The following procedures have been implemented in order to construct the TMTs analysis format.

Identifying the General Objectives of Each Course

English Language Curriculum Development and Updating Projects in Iraqi Universities is a general teaching guide which is recommended by the Ministry of Higher Education and Scientific Research to the Colleges of Education – Departments of English. The teaching guide includes the syllabus for each stage. The syllabi consist of many subjects. For the first stage (11) subjects, the second stage (10) subjects, the third stage (9) subjects, and the fourth stage (10) subjects.

The current study focuses on four courses that the researcher has chosen. Each course includes the course description, learning objectives (general objectives), course content and the required textbook.

Analyzing the Content of the Subjects Included in This Study

The content of each course included in this study has been analyzed. This involves the analysis of the following Books:

First stage/ Phonetics and Phonology: Better English Pronunciation by J. D. O'Connor.

Second stage/ Short Story: A Book of Short Stories by F. A. Razzak and A. Al-Muttalibi; and The Old Man and the Sea by E. Hemingway

Third stage/ Writing Academic Essay: Essay and Letter Writing by L. G. Alexander

Fourth stage/ Drama: Arms and the Man by George B. Shaw with three acts and no scenes; Waiting for Godot by S. Beckett with two acts and no scenes; and Death of the Salesman by A. Miller with two acts. The first act covers 12 scenes and the second act includes 14 scenes.

Deriving the Behavioral Objectives for Each Course

In order to construct the tests analysis format, a list of

Behavioral Objectives for the four courses has been designed. The list of the Behavioral Objectives consists of the Behavioral Objectives and their levels and according the cognitive domain of Bloom's taxonomy –the revised version, 2001.

Consequently, four lists have been developed. The following are brief description for each course and the number of the Behavioral Objectives according to the content:

Phonetics and Phonology

Phonetics and phonology are one of the subjects that is taught in the first stage at the Departments of English/ Colleges of Education. The researcher has designed a list of Behavioral Objectives for this course. The list consists of (108) of the Behavioral Objectives. However, all the Behavioral Objectives are designed

Short Story

Short Story is one of the subjects that is taught in the second stage at the Departments of English/ College of Education. After designing the list of objectives, (142) Behavioral Objectives have been formulated.

Writing Academic Essay

Writing Academic Essay is one of the subjects that is taught in the third stage at the Departments of English/ Colleges of Education. The list of the Behavioral Objectives that has been designed by the researcher consists of (238) Behavioral Objectives.

Drama

Drama is one of the subjects that is taught in the fourth stage at the Departments of English/ Colleges of Education. The list which has been formulated by the researcher includes the Behavioral Objectives for three plays. The first play, Arms and the Man includes (115) Behavioral Objectives. The second novel, Death of a Salesman involves (139) Behavioral Objectives. The third novel, waiting for Godot covers (30) Behavioral Objectives.

Adopting the List of the Action Verbs

The researcher has adopted Anderson and Krathwohl's (2001) list of action verbs with each behavioral objective. (See Appendix 1)

Collecting the Tests' Samples

The researcher has collected four samples of TMTs from the Department of English at the College of Education for Women.

The Tests Analysis Format

The tests analysis format has been constructed relying on the above procedures and as shown below:

University of Baghdad - College of Education for Woman Subject: Stage, Final Exam 2023- 2024.

The researcher has randomly selected a sample consisting of (50%) of the total TMTs for the college. Thus, two TMTs have been chosen to be analyzed by the researcher according to the test analysis format and to ascertain the reliability of the test analysis format. The TMTs which have been chosen are:

1. Phonetics and Phonology– first stage
2. Writing Academic Essay –third stage

Face Validity

McNamara (2000) refers to face validity as the extent to which an instrument meets the expectations of those involved in its use; e.g., administrators, teachers, candidates and test score users. Ebel (1972) states that face validity is the degree to which an instrument looks valid, and appears to measure the knowledge or abilities it claims to measure. It is used by asking experts to decide whether the tests analysis formats represent the features that we want to measure.

In order to ensure the face validity of the instrument, the lists of the Behavioral Objectives, the list of verbs associated with the objectives and the tests analysis formats, and the four analyses formats mentioned above have been exposed to a jury of 16 teaching staff members, well known for their long experience in the field of TEFL and they have agreed on the suitability and validity in 100%.

Reliability

Reliability, as validity, is another important characteristic of evaluation tools. It refers to the degree of stability of a measurement that is made reliability under different conditions or by different observers (Law, *et al.*, 1998).

Relying on the nature of the current study, the types of reliability that have been used are the intra-rater reliability and inter-rater reliability.

The reliability of the instrument is illustrated as follow:

Intra-rater reliability

The researcher has analyzed the two TMTs for the first time. After three months, the researcher has analyzed it for the second time using the same analysis tool.

Inter-rater reliability

The researcher has given the two tests analysis formats to a specialist. Then, she has clarified the aim of the study and its significance. Moreover, she has illustrated how to analyze the TMTs in order to follow individually the same procedure.

The researcher has collected the TMTs results in the form of frequencies and percentages in order to find the percentage of agreement between the researcher analysis and the other specialist.

Scott formula has been used to find out the percentage of agreement between the number of analysis times. However, the reliability of the test analysis format has been shown in Table 2.

Table 2: Test Analysis Formats’ Reliability

TMTs	Intra-rater reliability	Inter-rater reliability	Mean
College of Education for Women	0.91	0.82	0.85.5

Table (2) above shows that the reliability coefficient values are acceptable since they exceed (0.80) (Al-Said, 1979).

The Final Administration

The researcher has analyzed the TMTs’ samples according to the constructed format. (See Appendices 2, 3, 4 & 5).

The Mathematical and Statistical Means

The following mathematical means are used to analyze the data:

1. The percentage is used as a mathematical mean to calculate the percentages of objectives and in the TMT.
2. Scott formula is used as a mean to find out the inter-rater reliability:

$$\angle = \frac{P_o - P_e}{1 - P_o}$$

Where:

\angle = Inter-rater reliability

P_o = Percentage of observed agreement

P_e = Percentage of agreement expected by chance (Witcher, 2000).

RESULTS AND DISCUSSION

Results Related to the Aim

In order to evaluate Iraqi EFL University TMTs in light of the courses’ objectives, the percentages of the courses’ objectives and the percentages of the objectives in the TMTs have been calculated for the following courses:

- a. Phonetics and Phonology / first stage
- b. Short story / second stage
- c. Writing Academic Essay / third stage
- d. Drama / fourth stage

The Analysis of Teacher-Made Tests at the College of Education for Women

The TMTs of the following courses have been analyzed:

Phonetics and Phonology

The TMT in Phonetics and Phonology which is presented for the first stage students includes six questions with (83) items. The first question of essay type consists of seven items, the second question which is also of essay and supplying type includes six items, the third question is of short answer type involves thirty items, the fourth question which is of supplying type covers two items, the fifth question which is of short answer type encompasses ten items of three requirements with a total thirty items, while the sixth question which is also of short answer

type subsumes eight items.

After finding the percentages, the results reveal that the level of remembering yields (10) items with a percentage of 12%, understanding produces (33) items with 40%, applying possesses (40) items with 48%, while the levels of analyzing, evaluating and creating hold no items in the

test. The above percentages as compared with percentages of each level in the course objectives, which are 34% for remembering, 34% for understanding, 14% for applying, 13% for analyzing, and 5% for evaluating, show that the TMT does not achieve the percentages of the objectives included in the course (Table 3 and Figure 1).

Table 3: The Results of TMT Analysis in Phonetics and Phonology Course at the collage of Education for Woman

Level of cognitive behavioral objectives	No. of test items	Percentage of objectives in the TMT	Percentage of objective in the course
Remember	10	12%	34%
Understand	33	40%	34%
Apply	40	48%	14%
Analyze	-	-	13%
Evaluate	-	-	5%
Create	-	-	0%
Total	83	100%	100%

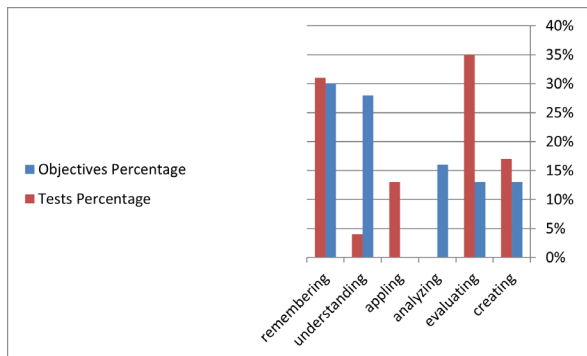


Figure 1: The Percentages of Objectives in Phonetics and Phonology Course and in the TMTs

Short Story

The TMT in short story which is presented for the second stage students covers six questions with (7) items.

The first, second, third, fourth and fifth questions which are all of essay type and each question includes one item, whereas the sixth question which is also of essay and supplying type encompasses two items.

After finding the percentages, the results submit that the level of understanding yields (1) item with a percentage of 14%, evaluating produces (5) items with 72%, creating possesses (1) item with 14%, whilst the levels of remembering, applying and analyzing hold no items in the test.

The above percentages as compared with percentages of each level in the course objectives, which are 33% for remembering, 28% for understanding, 0% for applying, 21% for analyzing, 14% for evaluating, and 4% for creating, show that the TMT does not represent the percentages of the objectives included in the course (Table 4 and figure 2).

Table 4: The Results of TMT Analysis in Short Story Course at the College of Education for Woman

Level of cognitive behavioral objectives	No. of test items	Percentage of objectives in the TMT	Percentage of objective in the course
Remember	-	-	33%
Understand	1	14%	28%
Apply	-	-	-
Analyze	-	-	21%
Evaluate	5	72%	14%
Create	1	14%	4%
Total	7	100%	100%

Writing Academic Essay

The TMT in writing academic essay which is presented for the third stage students includes six questions with (24) items. The first question which is of essay type contains three items, the second question which is also of essay type comprises four items, the third question is

of transformation and do as required type implicates six items, the fourth question which is of essay type covers two items, the fifth question which is of re-arrange type encompasses seven items and the sixth question which is of essay type subsumes (2) items.

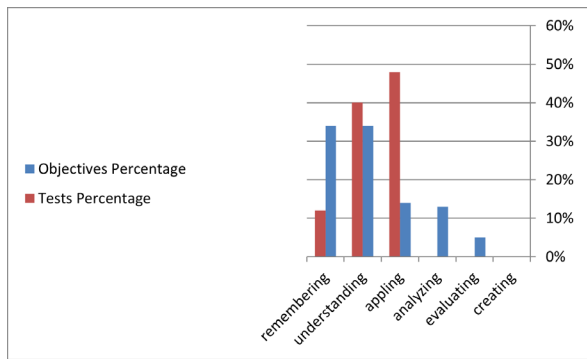


Figure 2: The Percentages of Objectives in Short Story Course and in the TMTs

After finding the percentages, the results reveal that the level of remembering yields (5) items with a percentage of 21%, understanding produces (10) items with 42%, applying possesses (7) items with 29%, creating results (2) items with 8%, while the levels of analyzing and evaluating hold no items in the test.

The above percentages as compared with percentages of each level in the course objectives, which are 34% for remembering, 28% for understanding, 8% for applying, 13% for analyzing, 12% for evaluating, and 5% for evaluating, show that the TMT does not achieve the percentages of the objectives included in the course (Table 5 and Figure 3).

Table 5: The Results of TMT Analysis in Writing Academic Essay Course at the College of Education for Woman

Level of cognitive behavioral objectives	No. of test items	Percentage of objectives in the TMT	Percentage of objective in the course
Remember	5	21%	34%
Understand	10	42%	28%
Apply	7	29%	8%
Analyze	-	-	13%
Evaluate	-	-	12%
Create	2	8%	5%
Total	24	100%	100%

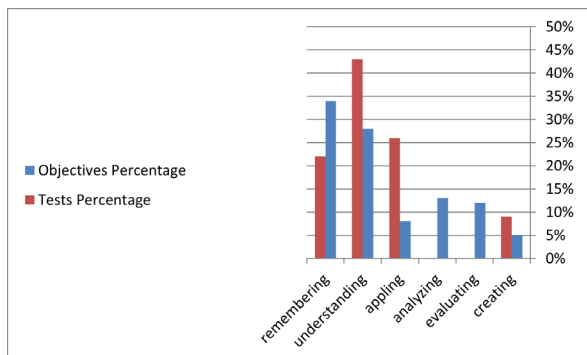


Figure 3: The Percentages of Objectives in Writing Academic Essay Course and in the TMTs

The third drama, waiting for Godot encompasses two questions which are of essay type. The results present that the level of analyzing produces (2) item with a percentage of 100%, while the other levels hold no items.

After finding the percentages of the objectives in the three plays, the results reveal that the level of analyzing produces (3) items with a percentage of 37.5%, evaluating possesses (3) items with 37.5%, and creating produces (2) items with 25%, while the levels of remembering, understanding and applying hold no items in the test.

The above percentages as compared with percentages of each level in the course objectives, which are 35% for remembering, 31% for understanding, 0% for applying, 18% for analyzing, 11% for evaluating, and 5% for creating, offer that the TMT does not exemplify the percentages of the objectives included in the course. See Table 6 and Figure 4.

The findings reveal that Iraqi EFL University TMTs at the College of Education for Women are not designed and developed according to the courses' objectives. It is also found that the TMTs items do not reflect the six cognitive levels of objectives. Moreover, it is found that most Iraqi EFL university teachers do not employ the action verbs adequately in their tests.

Drama

The TMT in drama which is offered for the fourth stage students includes eight questions with (8) items. The TMT covers three drama/ plays which are Death of a Salesman, Arms and the Man and Waiting for Godot.

The first drama, Death of a Salesman covers three questions which are of essay type. The results show that the level of analyzing produces (1) item with a percentage of 34%, evaluating yields (1) item with a percentage of 33%, and creating possesses (1) item with a percentage of 33%.

The second drama, Arms and the Man includes three questions which are also of essay type. The results reveals that the level of evaluating gets (2) items with a percentage of 67%, and creating possesses (1) item with a percentage of 33%.

Discussion of Results

The results show that Iraqi EFL university TMTs do not reflect the courses' objectives. This result is obtained because some of teachers may not have sufficient experience and are not trained in test construction skills.

Table 6: The Results of TMT Analysis in Drama Course at the College of Education for Woman

Level of cognitive behavioral objectives	1st Drama Death of a Salesman		2nd Drama Arms and the Man		3rd Drama Waiti -ng for Godot		Total		
	No. of test items	Percentage of objectives in the TMT	No. of test items	Percentage of objectives in the TMT	No. of test items	Percentage of objectives TMT	No. of test items	Percentage of objectives TMT	Percentage of objectives TMT
Remember	-	-	-	-	-	-	-	-	35%
Understand	-	-	-	-	-	-	-	-	31%
Apply	-	-	-	-	-	-	-	-	0%
Analyze	1	34%	-	-	2	100%	3	37.5%	18%
Evaluate	1	33%	2	67%	-	-	3	37%	11%
Create	1	33%	1	33%	-	-	2	25%	5%
Total	3	100%	3	100%	2	100%	8	100%	100%

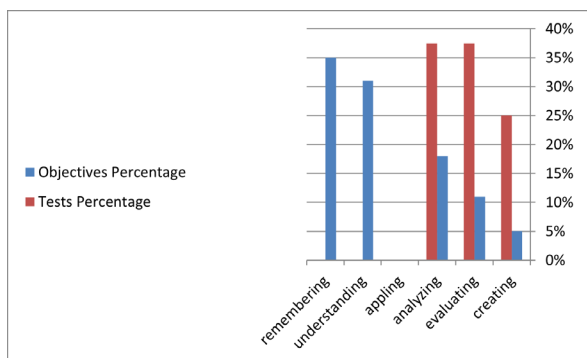


Figure 4: The Percentages of Objectives in Drama Course and in the TMT

They may lack the sufficient knowledge of the importance in developing the Behavioral Objectives in developing good language tests by describing the performance that should be measured. Without clearly stated objectives; teachers will have no clear idea about what exactly to test. A number of factors may lead to this result:

1. Some of the teachers are unable to use the suitable verbs in constructing their test.
2. Most of EFL teachers do not choose the suitable behaviors when they construct their tests.
3. Most of TMTs do not include the six levels of Blooms taxonomy –cognitive level.
4. The teachers do not use a table of specification when they construct their tests. Therefore, the TMTs' objectives percentages did not match the criteria of the courses objectives' percentages.

CONCLUSIONS

Based on the findings obtained from the present study and the researcher's observations, the following conclusions are drawn:

1. Despite the importance of the courses' objectives and specially the behavioral objectives in achieving the required outcomes, Iraqi EFL teachers do not pay enough attention to this important issue.
2. TMTs in the colleges of education concentrate on

some of Bloom's taxonomy levels –the cognitive domain and neglect the other levels.

3. Most of EFL teachers do not choose the suitable behaviors when they construct their tests.
4. Some of TMTs are constructed to focus on subjective items and ignore the objective items and vice versa.
5. The use of the action verbs in the test items has not been adequately employed as the analysis reveal.

Recommendations

In light of the results and conclusions of this study, the researcher forwards the following recommendations:

1. The behavioral objectives should be considered the initial step in designing the test; therefore, EFL teachers should design their tests as a reflection of what the students have studied during the course.
2. EFL teachers need to spend more time on writing their own tests and including various types of tests.
3. Iraqi EFL teachers need training in how to design their tests according to the criteria of courses' objectives.
4. Iraqi EFL teachers should take into consideration that the TMTs have to include the six levels of Bloom taxonomy –the cognitive domain.
5. The Ministry of Higher Education and Scientific research should provide an active testing guide which contains different types of tests that cover all the levels of Bloom's Taxonomy –cognitive domain.
6. EFL teachers have to recognize that the Behavioral Objectives describe the intended results of instructions and teaching, rather than the process of teaching itself.
7. A need to have assessment experts to review the test being done is important to ascertain the quality of testing.

REFERENCES

Bartlett, J. (2015). *Outstanding assessment for leaning in the classroom* (2nd ed.). New York: Author. https://www.routledge.com/Outstanding-Assessment-for-Learning-in-the-Classroom/Bartlett/p/book/9781138824508?srsId=AfmBOoonNLu-DWB5fT0DNtxERQaxKzCjPK5sPC6ehzOzZztQgR21_zdc.

Boud, D., & Falchikov, N. (2007). *Rethinking assessment in*

- higher education: *Learning for the longer term*. New York: Routledge. <https://www.routledge.com/Rethinking-Assessment-in-Higher-Education-Learning-for-the-Longer-Term/Boud-Falchikov>
- Board of Education for the City of Etobicoke. (1987). *Making the grade: Evaluating student progress*. Ontario: Prentice- Hall Canada.
- Boud, D. (1995). *Enhancing learning through self-assessment* (2nd ed.). London: Kogan Page.
- Brown D. H. (2001). *Teaching by principles* (2nd ed.). San Francisco: Pearson Education. <https://octovany.wordpress.com/wp-content/uploads/2013/12/ok-teaching-by-principles-h-douglas-brown.pdf>.
- Not mentioned. (2016) *English language curriculum development and updating project in Iraqi universities*. Ministry of Higher Education and Scientific Research: Department of Research and Development. https://ehs.uohamdaniya.edu.iq/english/wp-content/uploads/sites/16/2023/09/null-2_merged.pdf.
- Ebel, R. (1972). *Essentials of educational measurement* (2nd ed.). New Jersey: Hall Press.
- Fagan, L., & Spurrell, D. (1995). *Evaluating achievement of senior high school students in Canada*. Toronto: Canadian Education Association.
- Fulcher, G. (2010). *Practice language testing* (1st ed.). Abingdon: Routledge. https://www.routledge.com/Practical-Language-Testing/Fulcher/p/k/9780340984482?srsId=AfmBOoqydxzCVL9zwwfbd14Zv7jy_sRhLtgDOWxaUygFk5OhSOBbP_lx.
- Gay, L. R, Miller, G. E, & Anderson, P. (2012). *Educational research: Competences for analysis and application* (10th ed.). London: Pearson. https://yuli-elearning.com/pluginfile.php/4831/mod_resource/content/1/Gay-E%20Book%20Educational%20Research-2012.pdf.
- Genesee, F., & Upshur, J. (1996). *Classroom-based evaluation in second language education*. Cambridge: Cambridge University Press.
- Gorsuch G., & Griffiee, D. (2017). *Second language testing for student evaluation and classroom research*. Charlotte: Information Age. <https://www.amazon.com/Language-Testing-Evaluation-Classroom-Research/dp/1641130121>.
- Gronlund, N. (1999). *Measurement and evaluation in teaching*. New York: McGraw-Hill.
- Heaton, J. (1999). *Writing English language tests*. London: Longman. <https://octovany.wordpress.com/wp-content/uploads/2013/12/ok-writing-english-language-tests-j-b-heaton.pdf>.
- Magazine, W. H. (2018). *The importance of instructional objectives in teaching*. Retrieved From <http://wh-magazine.com/secondary-school/the-importance-of-instructional-objectives-in-teaching>.
- McNamara, T. (2000). *Language testing* (1st ed.). Oxford: Oxford University Press. <https://www.amazon.com/Language-Testing-Oxford-Introduction/dp/0194372227>
- Madsen, H. (1983). *Techniques in testing*. Oxford: Oxford University Press. <https://www.amazon.com/Techniques-Testing-Teaching-English-Language/dp/0194341321>.
- Nath, J., & Cohen, M. (2011). *Becoming an EC-6 teacher in Texas* (2nd ed.). Belmont: Wordsworth.
- Patel, R. N. (2011). *Educational evaluation: Theory and practice* (7th ed). New Delhi: Himalaya Publishing House. <https://himpub.com/product/educational-evaluation-theory-and-practice/>.
- Ramana, V. (2011). *Teaching large classes: Problems & prospects: Season your lectures with an active learning technique*. New Delhi: Education Publishing.
- Ramos, A. R., & Castillo, G. C. (2024). Best Practices and Challenges in the Assessment of Online Learning among Bsed Freshmen. *Journal of Tertiary Education and Learning*, 2(1), 33–39. <https://doi.org/10.54536/jtel.v2i1.2390>.
- Stobaugh, R. (2013). *Assessing critical thinking in middle and high schools: Meeting the common core* (1st ed.). Abingdon: Routledge. <https://www.routledge.com/Assessing-Critical-Thinking-in-Middle-and-High-Schools-Meeting-the-Common-Core/Stobaugh/p/book/9781596672338?srsId=AfmBOoMLsPJ4tCvfQVZjhS249renb-JbzkZdMXgBOzkfhfVl3VnGeb3>.
- Tuckman, B. W. (1975). *Measuring educational outcomes: Fundamentals of testing*. New York: Harcourt Brace Javanovich.
- Wiley, C. (2008). Traditional teacher tests in T. good (ed.), *21st century education: A reference handbook* (431-443). California: SAGE.
- Wilkinson, D., & Birmingham, P. (2003). *Using research instruments: A guide for researchers*. New York: Routledge. <https://thenigerianprofessionalaccountant.wordpress.com/wp-content/uploads/2013/04/using-research-instruments-a-toolkit-for-researchers.pdf>.
- Witcher, R. (2000). *After watergate: Nixon and the newsweeklies*. Maryland: University Press of America. <https://www.abebooks.com/9780761818557/After-Watergate-Nixon-Newsweeklies-Witcher-0761818553/plp>.