التحقق من مستوى التعقيد المعرفي لامتحانات الوزارة لطلبة الصف السادس الإعدادي العراقيين الذين يدرسون اللغة الإنجليزية كلغة أجنبية

م. م. عماد عبد الرضا عودة

المديرية العامة لتربية ديالى – العراق

mack.bride73@gmail.com

الملخص

تحاول هذه الورقة البحثية تقييم ودراسة مستوى التعقيد المعرفي لامتحانات الوزارة لطلبة الصف السادس الإعدادي العراقيين الدارسين اللغة الإنجليزية كلغة أجنبية. عينة الدراسة هي نماذج الامتحانات الوزارية النهائية للأعوام الدراسية (2019-2023) التي أُجريت على طلاب الفرع الأحيائي للدورين الأول والثاني. أداة الدراسة هي مقياس من إعداد الباحث. بعد استخراج الصدق الظاهري والثبات للأداة، تم وصف بيانات البحث وتحليلها باستخدام أساليب إحصائية مختلفة. توصلت الورقة البحثية إلى أن الاختبارات الوزارية قد أولت اهتمامًا كبيرًا للمستويات المعرفية للمعرفة والتطبيق نظرًا لسهولة صياغة أسئلة الاختبار وسهولة إجابة الطلاب عليها. ومع ذلك، أهملت صيغ الاختبارات ذات المستويات المعرفية العليا: التحليل والتقييم والابداع، وخاصة المستويين الأخيرين اللذين لم يحصلا على أي درجات في تحليل الدراسة. على الرغم من أن مهام الكتابة تهدف إلى قياس مستوى الأبداع، إلا أن غالبية الطلاب لا يزالون يحفظون نصوصًا جاهزة تتعلق بمواضيع الكتابة المحددة مسبقًا. قد يكون أحد الحلول هو زيادة مهام الكتابة وإلغاء تحديد المواضيع المطلوبة المُختبرة في الامتحانات النهائية. علاوة على ذلك، وُجد أن صيغ الامتحانات الوزارية لا تُحدد بنود اختبار تُعنى بالمهارات الشفهية: الاستماع والمحادثة، مع أن المنهج الدراسي يتضمن مهامًا تتعلق بكليهما. وأخيرًا، تُقدم الدراسة بعض التوصيات بناءً على النتائج.

الكلمات المفتاحية: تصنيف بلوم، المستوى المعرفي، مهمة الاختبار.

Investigating Cognitive Complexity Level of Iraqi EFL Sixth Grade Preparatory Learners' Ministerial Exams

Asst. Lect. Emad Abdul Ridha Oda

Ministry of Education- Diyala General Directorate - Iraq.

mack.bride73@gmail.com

Abstract

The research paper tries to rate and investigate cognitive complexity level of Iraqi EFL sixth grade preparatory learners' ministerial exams. The study sample is the forms of the final ministerial exams for the academic years (2019-2023) administered to the Biology branch students for both the first and the second sittings. The study instrument is a scale built by the researcher. After extracting face validity and reliability of the instrument, research data are described and analyzed by using different statistical techniques.

The research paper finds that the ministerial test formats have paid much attention to the cognitive levels of knowledge and application because of the ease with which test asks are formulated and the ease of answering them by students. However, the test formats have neglected higher order cognitive levels: analyzing, evaluating and creating, especially the last two levels which do not have any scores in the study analysis. Though writing tasks are meant to measure the level of creating, still the majority of the students memorize ready-made texts related to the pre-determined topics of writing. One solution might be to increase the writing tasks and to remove the ministerial specification of the topics tested in the final exams. Moreover, it is found that the ministerial exam formats do not specify test items that attend to oral skills: listening and speaking though the syllabus contains tasks related to both. Finally, the study puts forwards some recommendations based upon the findings.

Keywords: Bloom's taxonomy, cognitive level, test task.

Introduction

Statement of the Problem

Assessment is a major pillar of any educational process in which students' learning is measured by diverse procedures. Though these procedures signify enhancement, various issues related to learning assessment go on keeping unresolved (Veeravagu, Muthusamy, Marimuthu & Michael, 2010, p.205). Since the excellence of educational programs is based on the evaluation practice, exams play a significant role in learning for acting as one of the dimensions of evaluation.

While providing suitable exam questions at schools, composing the proper ones may be a problematic issue. In other words, choosing the right question is obviously the most difficult part of forming the exam paper, in addition to being the most time taking activity (Paul & Pawar, 2014, p.62). An exam paper is a traditional way of assessment- being the common choice of teachers evaluating the learners' degree of success in a particular lesson in which the necessary cognitive ability

of students is determined through the exam scores. That's to say, the questions presented on a paper determine whether the examination manages assessing the learners' performance or not. A good assessment requires an exam paper that covers different cognitive levels to accommodate diverse capabilities of learners (Jones, Harland, Reid & Bartlett, 2009).

Iraqi EFL learners college students are said to face difficulty or lag behind when they are exposed to tests tasks that measure high order thinking skills "Analyzing", and "Evaluating" and "creating" . However, they are familiar to deal with the questions that demand from them to do such cognitive tasks as define, choose, answer, match and recall. That's to say, they manage to deal only with tasks that measure lower thinking skills (Al-Khayyat, 2020, p.118)

Thinking skills learning is learning to employ facts, principles, and concepts in the thinking process. Thinking is defined as the act of withholding judgment in order to use past knowledge and experience to reach to new information, concepts and principles. Teaching students to think and be responsible is a major goal of education (Moore, 2007, pp. 119-120). Petty (2009, p.245) maintains that thinking skills refers to such processes

as analysis, problem solving, critical thinking, creative thinking, study skills, and the ability to create and express a coherent and well-justified argument. Knowledge is often a means to an end; it is only helpful when relevant to our thinking.

The human brain is capable of higher –order thinking, but we see so little of it in the normal course of student discussion and performance. Students are not thinking critically because we have not exposed them consistently to models or situations in school that require them to do so (Bruer, 1999: 5). Schooling, for the most part, demands little more than convergent thinking; its practices and testing focus on content acquisition through rote rehearsal, rather than the processes of thinking for analysis and synthesis. Too often, merely repeating the answer is considered more important than the process used to get the answer. Consequently, students and teachers have become accustomed to dealing with learning at the lowest levels of complexity (Sousa, 2001: 248-249).

In this regard, Petty (2009:245) asserts that weak teachers spend almost all their time teaching content; high order thinking skills are relatively ignored. Some teachers who follow an exclusively 'content focused' approach determine what content needs to be

taught and cram their teaching time with this. Yet, their students do poorly at assessment since it requires, besides content knowledge, high order thinking skills. These teachers fall into the content trap.

The best practice is when teaching develops thinking skills that facilitate mastery of the delivered content. Promoting thinking skills requires focusing more on reasoning tasks which pertain to higher level of thinking (level three upwards in Bloom's taxonomy) than on reproduction tasks which demand the student to repeat knowledge or skills that have been directly taught by the teacher or directly explained in resources. These tasks are lower on Bloom's taxonomy and they do not require the learner to process the material or to apply the learning (ibid: 13).

Giving the significance of thinking skills, this study is conducted to help specialists in test construction in the Iraqi Ministry of Education to have a better view of the current focus of the tests tasks with respect to cognitive complexity level. Also, the study may shed light on Iraqi EFL leaners' cognitive

1.1 AIMS

The current study aims at investigating which cognitive levels as

defined in Bloom's taxonomy are reflected in the Ministerial exam formats of the Iraqi EFL sixth grade preparatory schools students.

1.2 Limits

The current study is limited to the sixth grade preparatory English language exams for the academic years (2019-2023) administered to the Biology branch students for both the first and the second sittings except for Kurdistan region.

1.3 Value

Learning outcomes denote the students' accomplishments as an outcome from involvement in a certain set of teaching and learning activities. The three classifications of learning outcomes of students are: cognitive, affective and psychomotor (Bloom et al., 1956). It is essential that teaching and learning experiences, learning outcomes and the assessment are developed to comply with the three components (Biggs &Tang, 2011). Combination of these elements will determine conformity and uniformity within the syllabus where the expected learning outcomes align with the teaching, learning and assessment processes in a logical and consistent way.

Eventually, the success of any educational program depends on its evaluation system. Examinations are a

part of learning process which acts as an element in evaluation. The question tasks selection is vital for the test construction process. However, examination question paper composition is a multi-constraint optimization issue. Also, it is most significant and time-consuming activity. A number of constraints in this process are undervalued such as complexity and difficulty levels of the test tasks, total time duration for completion of the paper, and the total number of questions to be included in the test format (Paul, Naik & Pawar, 2014, p.61).

The findings of this paper is hoped to be insightful to Iraqi EFL preparatory school test constructors and designers in the Ministry of Education. They findings can give a better awareness with regard to the degree to which the exam formats cover the array of cognitive levels implied in the behavioral objectives of the syllabus. Also, the study results can be an invitation for teachers to reconsider the way they construct their monthly and mid-year exam formats through integrating tests tasks that tests different cognitive levels of thinking. Thereby teachers can enhance their students' preparedness to take the final examination.

Theoretical Background

2.1 Convergent Thinking Skills Vs Divergent Thinking Skills

Sousa (2001, p. 248) points out that cognitive psychologists' models of human thought have generally divided thought into two categories: convergent or lower-order thinking and divergent or higher-order thinking. A thinking model is good as long as it has the potential for encouraging higher-order thinking and the likelihood that teachers feel sufficiently comfortable with the model to make it a regular part of their classroom practice. Most models that describe the dimensions of thinking include the following four major areas:

1. Basic Processes. The tools we use to transform and evaluate information are:

-Observing: includes recognizing and recalling.

-Finding Patterns and Generalizing: includes classifying, comparing and contrasting, and identifying relevant and irrelevant information.

-Forming Conclusions Based on Patterns: includes hypothesizing, predicting, inferring, and applying.

-Assessing Conclusions Based on Observations: includes checking consistency, identifying biases and stereotypes, identifying unstated assumptions, recognizing over- and

under-generalizations, and confirming conclusions with facts.

 2. Domain-specific Knowledge. This refers to the knowledge in a particular Content area that one must possess in order to carry out the basic processes mentioned above.

3. Metacognition. This is the awareness one has of one's own thinking. It means that students should know when and why they are using the basic processes, and how these functions relate to the content they are learning (ibid:250). Metacognition requires students to self-monitor, self-assess and self-regulate their use of the study skills, setting themselves targets for experimentation and improvement (Petty: 2009:296).

4. Affective Domain. The affective domain processes (emotions and feelings) plays outstanding role in learning, and in the development and use of thinking skills. When students recognize the power of their own thinking, they use their skills more and solve problems for themselves rather than just waiting to be told the answers (ibid:250).

Bloom's taxonomy is the model that teachers can use to promote thinking in learning and teaching . The levels of the taxonomy are cumulative, that is, each level above "knowledge" includes all those of lesser complexity. A learner cannot comprehend material

without knowing it. Second, the lower three levels (knowledge, comprehension, and application) describe a convergent thinking process whereby the learner recalls and focuses what is known and comprehended to solve a problem through application. The upper three levels describe a divergent thinking process, because the learner's processing results in new insights and discoveries that were not part of the original information (Sousa, 2001, p.250).

Bloom's Taxonomy meets the four areas mentioned earlier that are included within most of the newer models describing the dimensions of thinking to a great extent. However, with regard to metacognition, it is implicit in the evaluation level. Since the model does not account for affective processes, the Taxonomy of the Affective Domain, developed by Krathwohl, Bloom and Masia (1964) can be referred to. It recognized the power of affect in attention and learning. Not only do we want students to learn cognitive information and skills and how to apply them but we also want them to appreciate and value their use. Developing positive attitudes in students toward learning enhance interest, increases retention, and should be a major goal of every teacher.

2.3 Bloom's Taxonomy and Revised Taxonomy

Boom's taxonomy has long been used by teachers everywhere to help plan lessons, to design curricula and assess learning outcomes. It is a functional assessment tool used to measure students' both learning and critical thinking skills according to the six stages of the taxonomy, is "a method of classifying educational objectives, educational experiences, learning processes, and evaluation questions and problems" (Paul, 1985 p. 39). While most of the assessments only cover calling up the memorized data. In such assessments, the questions simply refer to the first step of the taxonomy, though Bloom's taxonomy is composed of six steps in total: three steps in low order and three steps in high order cognitive skills (Eber & Parker, 2007). Bloom's Taxonomy addresses the arrangement of learning aims in the education process that educators appoint for learners. The cognitive domain within Bloom's taxonomy which is set to confirm a student's cognitive level (Haris & Omar, 2015) is the core of classifying statements according to what is expected from students to learn at the end of the instructional activities (Krathwohl, 2002). As an assessment practice, the employment of Bloom's Taxonomy yields important information for instructors. The taxonomy causes

instructors to be more conscious of the content and the process which they teach and assess, as well as indicating disparities between what is taught and what is assessed. Further, it can perform as a guide to evolve and expand the learning and assessment activities by supplying a concrete consciousness of the content and process— an instructor defines as essential in the development of learners' cognition (Kastberg, 2003, p. 405).

While the American Heritage Dictionary of the English Language (2016) refers to cognition as "the mental process of knowing, including aspects such as awareness, perception, reasoning, and judgment, or that which comes to be known, as through perception, reasoning, or intuition; knowledge", the cognitive domain addresses the knowledge and development of intellectual skills (Bloom, 1956). As we have already mentioned, this domain is grouped under six subsequent thinking levels: the first three levels which refer to the lower order thinking skills include remembering, understanding, and applying, while the next three levels refer to the higher-order thinking skills that contain analyzing, evaluating, and creating (Orey, 2010). Bloom's taxonomy is hierarchical and each step is presented at the upper steps as well. So, we can see the lower level at the higher level as well. For instance,

someone in the analyzing step can also function in remembering, understanding, and applying steps (Konza, 2011).

Since the 1950's, Bloom's Taxonomy has been used to structure the thinking process in education. Later research supported the concept that the natural thinking process begins with the lower levels of the Taxonomy, and proceeds to the higher levels. Yet, subsequent research revealed that up to 90 percent of teaching occurs at the knowledge level, which is the lowest of Bloom's six levels (Davidson & Decker, 2006).

Bloom's Taxonomy plays an important part in improving critical thinking of students; this is important as students need skills that will help them to solve problems and think critically, both in and out of the classroom. Armstrong (2015) provided detailed background information about this theory; specifically, Benjamin Bloom, with collaboration from Max Englehart, Edward Furst, Walter Hill and David Krathwohl, published a framework in 1956 to categorise educational goals of taxonomy of educational objectives, this became known as Bloom's Taxonomy. The framework focuses on six main categories, namely: knowledge, comprehension, application, analysis, synthesis, and evaluation and knowledge that are

necessary to implement skills and abilities into action.

All of the categories, with the exception of "knowledge" became known as the "skills and abilities" category. Each taxonomy is popularly remembered due to its categorisation which can be further subdivided from simple to concrete divisions. Armstrong (2015) explained these main categories as knowledge involving the recall of specifics and universals, the recall of methods and processes, or the recall of a pattern, structure, or setting (see figure 1). Each category will now be addressed in turn:

• Comprehension: refers to understanding or interpreting what has been communicated to ultimately make use of ideas and materials without relating them to some other material or idea. It is known as the ability of the mind to understand and perceive.

• Application: refers to the special use of an abstract to which something is put.
• Analysis: represents understanding the nature of something by breaking it into
constituent elements or parts to ensure that the idea interpreted is clear and the relation between the ideas is addressed properly.

• Synthesis: combines two or more elements together so that they take the form of a new element.m

• Evaluation: involves judging the merit, worth and significance by a set of standards.

There are several benefits from using the New Taxonomy, in particular it also supports the thinking level of the teachers and the students in the following ways:

1. Objectives or learning goals are important as it helps to establish the purpose of the educational interchange for both students and teachers.

2. Teachers can benefit from using the framework by helping them to establish objectives, it will also help them to clarify things that need be to be done by the students, as well as the teachers.

3. Having an organized set of objectives helps teachers to:

• Plan and deliver proper instructions to the students;

• Valid designing of assessment tasks and strategies; and,

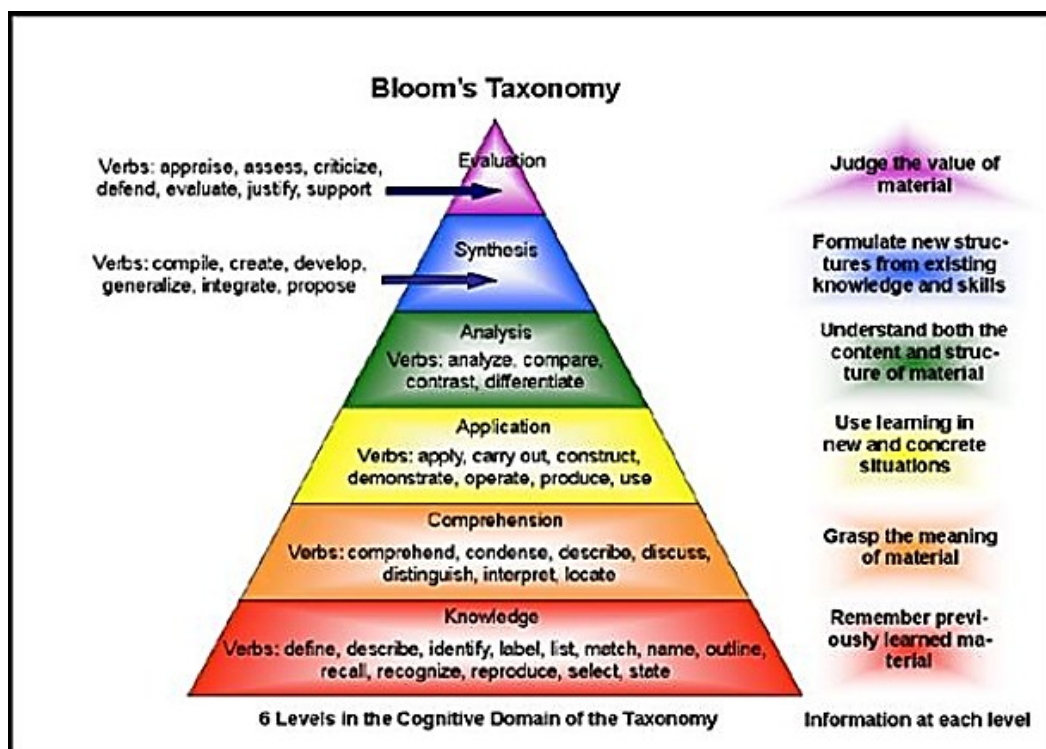• Ensuring instructions and assignments are designed along with the objectives (Armstrong, 2015).

Figure (1) Levels of Cognitive Domain (Lister, 2012, p.12)

2.4 Cognitive Complexity or Difficulty

One of the misinterpretations of two terms related to the use of Bloom's taxonomy that limits the use of the taxonomy to enhance the thinking of all students is that of difficulty and complexity. Complexity denotes the cognitive level in the taxonomy, that is to say what cognitive process the brain use to attend to the information or the learning task. While difficulty refers to the amount of effort that the learner must exert within a level of complexity to attain a learning objective. A learning activity can be increasingly difficult without becoming more complex. For example, the question "What is the capital of Rhode Island?" is at the knowledge level; and the question "What are the capitals of the state unions?" is still at the knowledge level, but more difficult. Recognizing the difference between these two concepts, the teacher can gain valuable insight into the connection between the taxonomy and student ability (Sousa, 2001: 258-259).

To move up Bloom's taxonomy with all students, teachers need to review the curriculum and remove the topics of least importance to gain the time needed for practice at the higher levels. An effective method for doing this is to set priorities among all the concepts in a curriculum, delete the

least important bottom 20 to 25 percent, and use the time gained by this sorting and paring to move all students up the taxonomy. Finally, they should take advantage of the power of positive transfer by integrating these concepts with previously taught material and connecting them to appropriate concepts in other curriculum areas (Ibid: 262).

What current research recommends is to recognize the limitations of tasks complexity and amount, rewrite curriculum, retrain teachers, and to encourage students to use their innate thinking abilities to process learning at higher levels of complexity. In other words, teachers need to teach their students how to organize content in such a way that it facilitates and promotes higher-order thinking. Teachers who believe they are teaching students how to think perpetuate the teaching-learning scheme that treats students as "vessels to be filled." However, accepting that we are sharpening their skills to facilitate thinking places the teacher in the proper position of guiding students through the more effective use of their innate processes and abilities. The teacher's role shifts, as a current saying goes, from being the "sage on stage" to the "guide on the side.' (Sousa, 2001: 248-49).

2. 4 Previous Studies

As far as the Iraqi EFL context is concerned, the literature on investigating the level of complexity of final exams is scarce. It seems not much has been done or published on this issue. One study, if any, is conducted by Al-Khayyat (2020). The study aims at investigating the revised Bloom's cognitive level (verbs) that are included in the Baccalaureate English Language exams. The sample of the study consisted of the Baccalaureate English Language exams' forms for the years 2016,2017,2018, and 2019 in the first and second attempts. Towards this end, the following null hypothesis has been set "there was no statistical differences in the percentage of using Bloom's cognitive taxonomy in the Baccalaureate exams that can be attributed to the cognitive levels". To achieve the aim and to prove the hypothesis of the study, the researcher built a scale and elicited its reliability and validity. The results showed that the Baccalaureate exams measure students' cognitive levels in "Remember", "Understand" and "Create" and have neglected students' cognitive levels in "Apply", "Analyze", and "Evaluate". However, the neglected cognitive levels are included in the "English for Iraq 6th preparatory 'Teacher Book' Guide". Also, the Speaking and listening skills are neglected too.

Regionally speaking, Al-Skaf 's (2017) study aimed at classifying the levels of assessment questions according to Bloom's taxonomy in English course for eleventh grade in Syria and to analyze the percentage of questions in each level (recall, comprehension, application, analysis, synthesis, and evaluation). The sample of the study consisted of student's book for eleventh grade in Syrian Arab Republic. The study followed the quantitative approach by making use of such tool as: an analysis card to measure the frequency of assessment questions. The results showed that assessment questions in English student's book for eleventh grade covered all levels of Bloom"s Taxonomy. The percentages were: recall 51.65%, comprehension 14.89%, application 9,02%, analysis 2.8%, synthesis 7.48%, and evaluation 2.63%. Thus, recall which is the lowest thinking skill was the most frequent level of questions whereas evaluation was the least occurring level.

METHODOLOGY

3.1. Sample of the Study

The current study sample includes the final ministerial exams for the academic years (2019-2023) administered to the Biology branch students for both the first and the second sittings.

3.2 The Instrument of the study

To achieve the study aim, the researcher built a scale that has a percentage table measuring the inclusion of Bloom's Taxonomy levels in the Baccalaureate English language exams according to the academic year. The scale is made up of two main dimensions: The Bloom's cognitive level (the new version 2001) and the frequency of the level as reflected in the individual test tasks of each year (See Table **1**).

Table (1): The scale of the study

Cognitive complexity level	Frequency	Year
Remembering Define, Identify, List, Name, Recall, Recognize, Re-Peat, Re-write		
understanding Choose, Answer, Describe, Match, Differentiate, between, Express, give in own word, Discuss, Complete.		
Applying Practice, Use, Relate, Interpret, Apply, Interact, demonstrate, show + …		
Analyzing Analyze, Appraise, compare, contrast, Develop, Contrast, Develop, Diagram, Distinguish, Draw, Evaluate, Infer, Question, Predict, recognize		
Evaluating Compare, Critique, Evaluate, Judge, Measure, Select, Predict, Test, Score, Assess, Summarize.		
Creating Arrange, Collect, Write, Design, Organize, Compose, Propose, Set-Up Modify, Assemble	.	

3.2.1 Face Validity

Validity is defined as the extent to which the test measures what it is intended to measure. It is the most significant principle of language assessment (Brown, 2004, p.22).

Face validity "Is the most basic kind of validity; it is a judgment by the scientific community that the indicator really measures the construct" (Riazi, 1999:96).

As mentioned above, the scale is built by the researcher who has exposed it to jury members in ELT to judge on its face validity and suitability to the study aim. And they found that it has face validity and is suitable after making some modifications and giving recommendations.

3.2.2 Reliability of the Scale

Reliability means accuracy and consistency in the performance of individuals and the stability of results over time. A reliable scale gives the same results if it is applied to the same individuals again. Also, an individual will obtain the same scores if she is exposed to the same tool and under the same conditions.

To ensure the reliability of the tool, the researcher conducted the following:

1. Analyzing ministerial questions by herself for the first time, then analyzing them for the second time after a period of time, using the same analysis tool that was used the first time.

2. The researcher drew a random sample consisting of (20%) of the total questions of all years and presented it to another analyst who possesses the

necessary qualifications to carry out the analysis process according to its correct methodology. She was given the analysis scale and was able to recognize the study aim and importance. Then, she was introduced to the method of analysis, its rules, constraints, and procedures. Later, she was given the opportunity to do the analysis process alone.

3. Collecting the results of the analysis conducted by the researcher and her colleague in the form of frequencies and percentages in preparation for calculating the percentage of agreement between them.

4. Extracting the percentage of agreement between the two times analyses using Scott's equation, and the results are as shown in Table (2):

Table (2): the analysis invariance coefficient for ministerial questions.

The test tasks	Intra-rater reliability	Inter-rater	mean
The ministerial questions	92%	88%	90%

DATA ANALYSIS, DISCUSSIONS OF RESULTS, CONCLUSIONS AND RECOMMENDATIONS

In this part of the research, the researcher presents the findings related to the evaluation of ministerial questions of the English language according to Bloom's cognitive levels. She provides a discussion of these results and sets up mechanisms for their correct implementation based on specific and planned scientific foundations.

The results can be presented and interpreted as follows:

The study aim is which cognitive levels as defined in Bloom's taxonomy are

reflected in the Ministerial exam formats of the Iraqi EFL sixth grade preparatory schools students.

Towards this goal, the frequencies and percentages of ministerial questions are calculated for each year separately as follows:

First: Ministerial Questions for the Year (2019):

The frequencies and percentages of ministerial questions for the year (2019) are calculated for each level of knowledge. The total number of questions for this year is (108). The results are shown in Table (3) and Figure (2).

Table (3): Frequencies and percentages of ministerial questions for the year (2019) according to Bloom's classification

No.	Cognitive level	Frequency	percentage
1	Knowledge	34	%31
2	Remembering	30	%28
3	Application	32	%30
4	Analysis	12	%11
5	Evaluation	0	%0
6	Creating	0	%0
	Total	108	%100

مجلة اليرموك العدد الأول المجلد 23 لسنة 2025
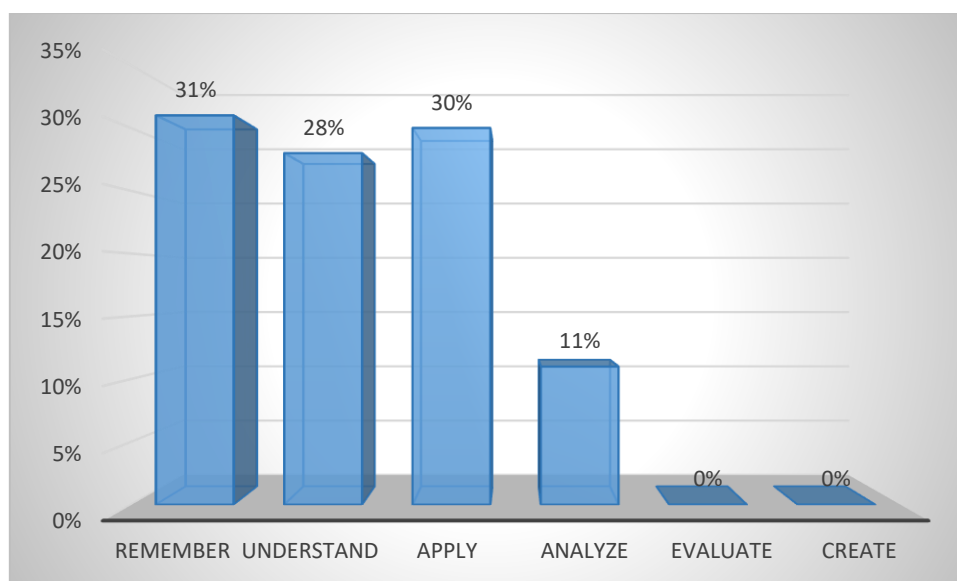
Figure (2) :shows the percentages of knowledge levels for ministerial questions for the year (2019)

It is clear from the above table and figure that the total ministerial test tasks for the English language course for the year (2019) amounted to (108) questions. The level of knowledge obtained (34) frequencies and percentage of (31%). The level of understanding got (30) frequencies and percentage of (28%). While the frequencies of the level of application was (32) the percentage was (30%). The level of analysis got (12) frequencies 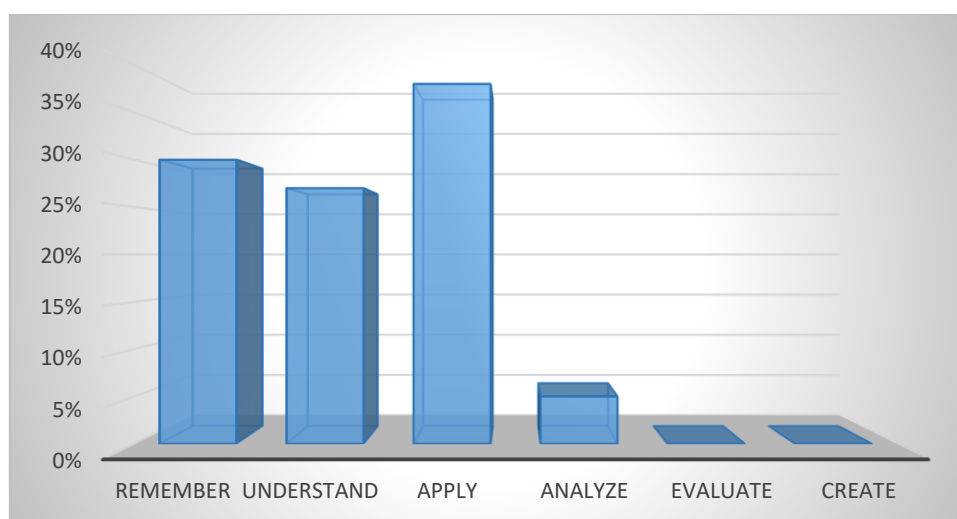and a percentage of (11%). As for the levels of evaluation and creation, they did not obtain any frequencies in the ministerial questions for the year (2019). Second: Ministerial Questions for the year (2020):

The frequencies and percentages of ministerial questions for the year (2020) were calculated for each level of knowledge. The total number of questions for this year was (110). The results were as shown in Table (4) and Figure (3).

Table (4): Frequencies and percentages of ministerial questions for the year (2020) according to Bloom's classification

No.	Cognitive level	Frequency	percentage
1	Knowledge	33	%30
2	Remembering	30	%27
3	Application	42	%38

262

4	Analysis	5	%5
5	Evaluation	0	%0
6	Creating	0	%0
Total		110	%100

Figure(3) :shows the percentages of knowledge levels for ministerial questions for the year (2020)

From the above table and figure, it becomes obvious that the total ministerial test tasks for the English language subject for the academic year (2020) were (110) questions. The level of knowledge scored (33) frequencies and percentage of (30%). While the level of understanding got (30) frequencies and its percentage was (27%). The frequencies of level of application were (42) and its percentage was (38%). However, the level of analysis was reflected in (5) frequenci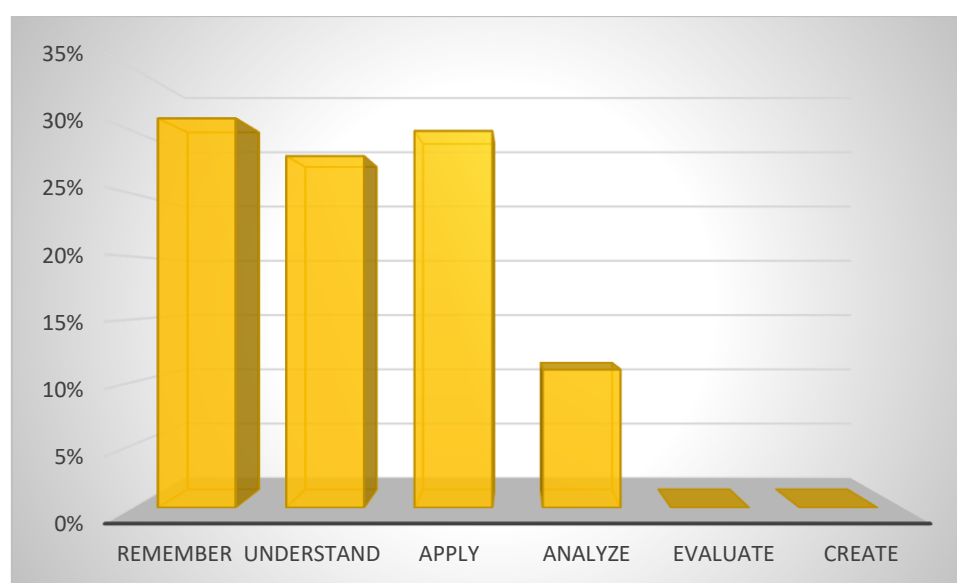es (5%) of percentage. As for the levels of evaluation and creation, they did not obtain any scores in the ministerial questions for the year (2020).

Third: Ministerial Questions for the Year (2021):

The frequencies and percentages of ministerial questions for the year (2021) were calculated for each level of knowledge. The total number of the test tasks for this year was (108). The results were as shown in Table (5) and Figure (4).

Table (5): Frequencies and percentages of ministerial questions for the year (2021) according to Bloom's classification

No.	Cognitive level	Frequency	percentage
1	Knowledge	33	%30
2	Remembering	29	%27
3	Application	40	%38
4	Analysis	6	%5
5	Evaluation	0	%0
6	Creating	0	%0
Total		108	%100

Figure(4): shows the percentages of knowledge levels for ministerial questions for the year (2021)

It is clear from the above table and figure that the total test tasks for the year (2021) was (108) questions. The level of knowledge obtained (33) frequencies and (31%). Whereas the level of understanding got (29) frequencies and (27%). The frequencies of the level of application were (40) and the percentage was (37%). The level of analysis scored (6) frequencies and (5%). However, the levels of evaluation and creation did not score any

<table>
</table>

The frequencies and percentages of ministerial questions for the year (2022) were calculated for each knowledge level. The total number of questions for this year was (108). The results were as shown in Table (6) and Figure (5).

frequencies in the ministerial questions for the year (2021).

Fourth: Ministerial Questions for the Year (2022):

Table (6): Frequencies and percentages of ministerial questions for the year (2022) according to Bloom's classification

No.	Cognitive level	Frequency	percentage
1	Knowledge	32	%30
2	Remembering	28	%26
3	Application	44	%41
4	Analysis	4	%3
5	Evaluation	0	%0
6	Creating	0	%0
Total		108	%100
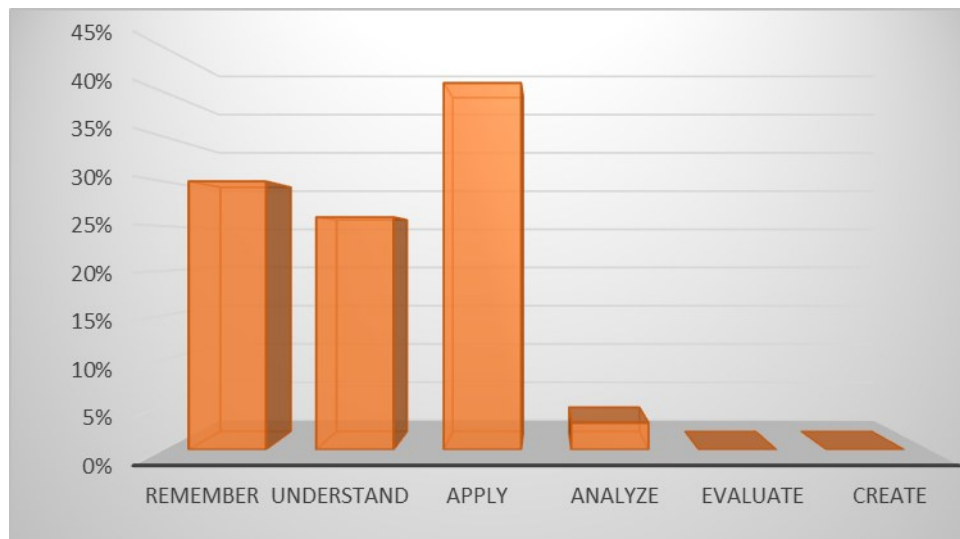
Figure (5): shows the percentages of knowledge levels for ministerial questions for the year (2022)

Table (6) and figure (5) illustrate that the total tasks of the ministerial exam format for the year (2022) amounted to (108) questions. The level of knowledge obtained (32) frequencies and (30%). The level of understanding got (28) frequencies and (26%). While the level of application scored (44) frequencies and (41%). The frequencies of the level of analysis were (4) and the percentage was (3%). As for the levels of evaluation and creation, they did not obtain any score in the ministerial questions for the year (2022).
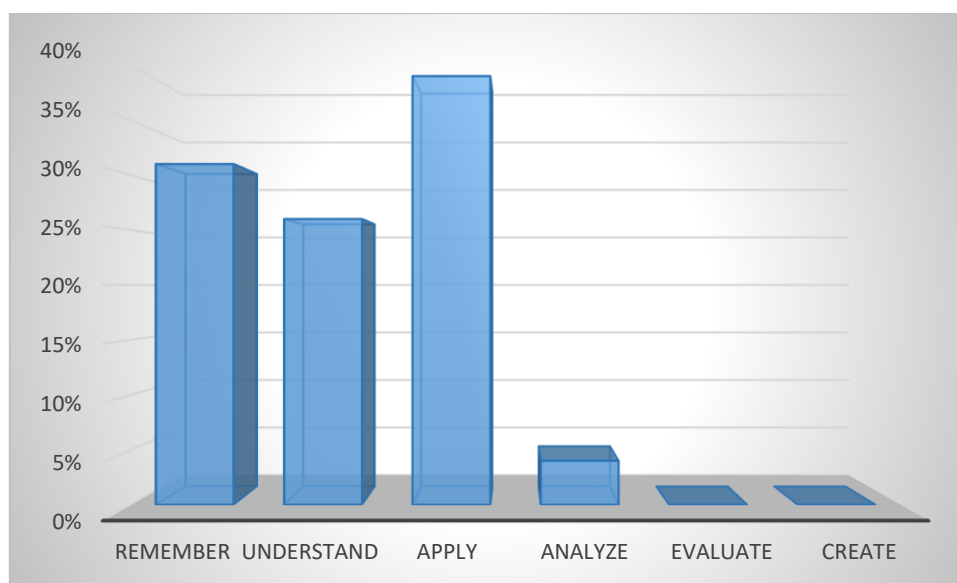
Fifth: Ministerial Questions for the Year (2023):

The frequencies and percentages of the test tasks of the ministerial questions for the year (2023) were calculated for each cognitive level. They reached the total number of (108) tasks for this. The results were as shown in Table (7) and Figure (6)

Table (7): Frequencies and percentages of ministerial questions for the year (2023) according to Bloom's classification

No.	Cognitive level	Frequency	percentage
1	Knowledge	33	%31
2	Remembering	28	%26
3	Application	43	%39
4	Analysis	4	%4
5	Evaluation	0	%0
6	Creating	0	%0
Total		108	%100

Figure (6) shows the percentages of knowledge levels for ministerial questions for the year (2023)

It is clear from the above table and figure that the total test items for the English language course for the year (2023) amounted to (108) items. The level of knowledge obtained (33) frequencies and (31%). Moreover, the level of understanding got (28) frequencies and (26%). While the level of application was reflected in (42) frequencies and (39%). Wheras the level of analysis scored (4) frequencies and (4%). As for the levels of evaluation and creation, they did not obtain any score in the ministerial question format for the academic year (2023).

Discussion of Results

The present study examined the cognitive complexity levels reflected in Iraqi EFL sixth grade preparatory learners' ministerial exams along a five-

year period. As the analysis uncovered, the majority of test tasks are heavily focused on lower-order cognitive levels—specifically, knowledge (remembering) and understanding—with a substantial number of items addressing the application level. However, higher-order cognitive skills, particularly analysis, evaluation, and creation, are seen to be either minimally represented or entirely absent.

Interestingly, previous studies in the region align with these findings. For instance, Al-Khayyat (2020) reported that Baccalaureate English Language exams in Iraq predominantly measure lower cognitive levels while neglecting critical higher-order skills. Similarly, Al-Skaf (2017) noted that in Syrian secondary schools, assessment questions target recall and comprehension to a greater extent and

evaluation and synthesis to a much lesser context. This consistency suggests a regional trend in test construction that emphasizes ease of question formulation and scoring over the development of critical thinking and creative problem-solving abilities.

Moreover, this study results support the notion that the prevailing test formats are influenced by practical constraints. The ease with which questions targeting lower cognitive skills can be written and answered makes them more attractive for examiners. However, this focus can lead to a narrow assessment of students' capacities, failing to capture the full spectrum of cognitive processing that modern educational standards aspire to promote. As Sousa (2001) argues, an overemphasis on rote learning limits the opportunity for learners to engage in divergent thinking and deeper cognitive processing.

It is particularly worrying to have this state of negligible presence of tasks measuring analysis, evaluation, and creation given the essential role of these skills in preparing students for real-world challenges. Without opportunities to engage in tasks that demand critical analysis or creative synthesis, learners may be underprepared for higher education and future professional contexts that require robust problem-solving abilities. In this light, the current study echoes earlier calls for reform in test design and curriculum alignment, urging educators to integrate higher-order tasks into exam formats.

Accordingly, both the current study and previous research (e.g., Al-Khayyat, 2020; Al-Skaf, 2017) converge on the finding that EFL assessment in the region is predominantly skewed toward lower cognitive skills. This recurring pattern underlines the need for test constructors and educational policymakers to revisit and revise assessment practices. Expanding the range of cognitive skills tested—particularly those at higher levels—could foster a more balanced and comprehensive evaluation of student learning outcomes, ultimately contributing to improved teaching and learning in the EFL context.

Conclusions and Recommendations

From the current results, the researcher concludes the following:

1. The noticeable increase in the level of knowledge test items, as it obtained a percentage (31%) for the year (2019), a percentage of (30%) for the year (2020), (31) for the year (2021), and a percentage of (30) for the year (2022), and a percentage of (31) for the year (2023). These percentages are high, and therefore it can be said that EFL teachers pay good attention to the level of knowledge or remembering because

of the ease with which test asks are formulated and the ease of answering them by students. The researcher believes that the value of knowledge can be downgraded if it is not employed, applied, and benefited from in new learning situations. Excessive concern for remembering information leads to transforming teaching into mere memorization and recitation or retelling of information learned.

2. The level of understanding obtained a percentage of (28%) for the academic year (2019), (27%) for the year (2020), (27%) for the year (2021), (26%) for the year (2022), and a percentage of (26%) for the year (2023). The researcher believes that these percentages are acceptable to some extent since the level of understanding is an important level and must be represented in the questions.

3. The percentage of questions that measure the level of application is high, as it obtained a percentage of (30%) for (2019), (38%) for (2020), (37%) for (2021), (41%) for (2022), and (39%) for (2023). On one hand, the researcher believes that this percentage is too high and exaggerated, as the test items that represent the level of application were drawn at the expense of other cognitive levels.

On the other hand, test constructors should take into account, in designing tasks that measure the

level of application, that these tasks do not remain static and stagnant in nature, taking a single pattern. But rather they should be renewable and changeable and that their authors introduce new techniques that gauge the application level.

4. The percentages of the test items measuring the level of analysis are too low as they received the following percentages: (11%) for the year (2019), (5%) for the year (2020), (5%) for the year (2021), (3%) for the year (2022), and (4%) for the year (2023). When compared to importance of this cognitive level and to the role it plays in the development of EFL students' performance in the subject.

5. There is no percentage in the ministerial questions that measure higher cognitive abilities and levels of evaluating and creating in ministerial questions for all years. The researcher attributes this to the fact that the test designers may put these questions without taking into account these Levels. Perhaps one of the important reasons behind the neglect of higher levels of knowledge is the failure of those in charge of tests construction to have a clear strategy based on specific learning and teaching objectives, a strategy that can take care and attention to measure these important levels.

In light of the following conclusions, the researcher recommends the following:

1. Iraqi EFL final exam constructors need to include test items that measure higher order thinking skills (analyzing, evaluating and creating).

2. Iraqi Ministry of Education should train and inform EFL preparatory school teachers on the standardized procedures followed in test constructions and design such as using the test map in the most appropriate way to make their exam formats.

3. Iraqi Ministry of Education need to educate EFL preparatory school teachers on the significant relationship between behavioral objectives set in the Teachers' Guide and the process of assessment and evaluation of the teaching- learning process.

4. Test makers at the Iraqi Ministry of Education need to increase the writing tasks in the syllabus and to remove the ministerial specification of the topics tested in the final exams.

5. Moreover, the ministerial exam formats must specify test items that attend to oral skills: listening and speaking since the syllabus contains tasks related to both

6. Iraqi Ministry of Education has to construct a standardized test bank that serve as a reference point for test designers, EFL teachers and learners. It can encourage the researchers to present research papers on this topic.

7. Iraqi Ministry of Education needs to train and educate EFL teachers on the importance of divergent thinking skills and on the best tactics that they can employ towards applying this objective in their classes.

References

[1] Al-Khayyat, A. (2020). Investigating the Level of Bloom's Cognitive Verbs Included in Baccalaureate English Language Exams in Iraq. Journal of University of Shanghai for Science and Technology, 22(11), 118–129.

[2] Al-Skaf, M. M. (2017). Maysoun shams aldiyn. An analytic study of the levels of evaluation questions in English courses in Syrian secondary schools according to Bloom's Taxonomy. AL-USTATH Number extension 222–Volume One, 1438 AH.

[3] Armstrong, P. (2015). Bloom's Taxonomy. Centre for Teaching: Vanderbilt University. [Online]. Available at: http://cft.vanderbilt.edu/guides-sub-pages/bloomstaxonomy/ [Accessed on: 25 Aug 2015].

[4] Biggs, J., & Tang, C. (2011). Teaching for quality learning at university (4th ed.). McGraw-Hill Education.

[5] Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (1956). Taxonomy of educational objectives: The classification of educational goals. Handbook I: Cognitive domain. David McKay Company.

[6] Bruer, J. (1999). The myth of the first three years. New York: Free Press. [Also reprinted in International Journal of ICT Research and Development in Africa, 4(1), 61–75, January–June 2014].

[7] Davidson, K., & Decker, T. (2006). Bloom's and beyond: Higher level questions and activities for the creative classroom. Marion, IL: Pieces of Learning. Retrieved December 9, 2009, from http://www.mysdcc.sdccd.edu/Importance_of_a_Lesson_Plan.htm.

[8] Eber, R., & Parker, T. (2007). Enhancing cognitive skills through Bloom's taxonomy: An exploratory study. Journal of Educational Research, 100(2), 150–164.

[9] Haris, A., & Omar, M. (2015). Evaluating cognitive domains in language assessment: Applying Bloom's taxonomy. International Journal of Language Education, 7(1), 37–49.

[10] Jones, K. O., Harland, J., Reid, J., & Bartlett, R. (2009). Relationship between examination questions and Bloom's taxonomy. In Frontiers in Education Conference, 2009. FIE'09. 39th IEEE, 1–6.

[11] Kastberg, S. (2003). The role of assessment in developing higher-order thinking skills. Educational Measurement, 40(5), 403–410.

[12] Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: An overview. Theory into Practice, 41(4), 212–218.

[13] Lister, R. (2012). The CC2013 Strawman and Bloom's Taxonomy. ACM Inroads, 3(2), 12–13.

[14] Moore, T. J. (2011). Critical Thinking and Language: The Challenge of Generic Skills and Disciplinary Discourses. Continuum International Publishing Group.

[15] Paul, D. V., Naik, S. B., & Pawar, J. D. (2014). An Evolutionary Approach for Question Selection from a Question Bank: A Case Study. International Journal of ICT Research and Development in Africa (IJICTRDA), 4(1), 61–65.

[16] Paul, R. W. (1985). "Bloom's taxonomy and critical thinking instruction." Educational Leadership, 42, 36. Association for Supervision & Curriculum Development.

[17] Petty, G. (2009). Evidence-Based Teaching: A Practical Approach. Nelson Thornes Ltd.

[18] Sousa, D. A. (2001). How the brain learns. Corwin Press.

[19] Veeravagu, J., Muthusamy, C., Marimuthu, R., & Michael, A. S. (2010). Using Bloom's Taxonomy to Gauge Students' Reading Comprehension Performance/Utiliser La Taxonomie De Bloom Pour Evaluer Les Performances De Comprehension Ecrite Des Eleves. Canadian Social Science, 6(3), 205–222.