

## Assessing EFL University Learners' Writing Performance

RAHMOUNA ZIDANE

*Tlemcen Üniversitesi Edebiyat Fakóltesi*

**Abstract:** The assessment of learners of English as a foreign language (EFL) relies on a set of procedures and methods. However, it may be characterized by certain difficulties in designing and scoring examination questions. The following article provides information about testing EFL university students by focusing mainly on the learners' writing performance. To study the factors that affect the tests' results, a set of examination topics and a sample of learners' scores were analysed in order to highlight the issue of the topic effect and its position within the process of assessment. The results obtained led to the conclusion that the subject of the examination plays a role in determining the quality of the learners' achievement.

**Keywords:** Testing, Topic Effect, Performance.

The aim of teaching English as a foreign language (EFL) to Algerian university students is to make the learners master language to get a degree in order to become teachers of English or to use this language for other occupational purposes. The assessment of EFL learners tries to provide information about the students' ability to speak fluently and to generate accurate written productions. This process may incorporate different methods including analytic and holistic scoring. Thus, students are tested through oral or written examinations depending on the taught unit. However, the majority of the modules rely on written examinations that often involve the production of essays.

When testing EFL learners, the teachers may face certain difficulties in constructing examination questions and scoring the testees answers. Thus, it is necessary to make research about the factors that may create problems for the teachers and the learners when dealing with examinations. The present article attempts to shed light on the assessment of EFL learners. In fact, the purpose of this study is to diagnose the EFL teaching and testing situation for the sake of developing the teaching programmes and the testing methods.

### **Testing EFL Undergraduate Students**

In Algerian universities, English is taught as a foreign language. At the institutes of science and technology, this language is taught for specific purposes (ESP) to those students who are learning scientific subjects and need English as a medium of instruction to have access to special references provided in this language. On the other hand, at the institutes of languages and arts, English is considered as a language of study. Thus, all the four skills (speaking, listening, reading and writing) are taken into account in addition to content areas disciplines (phonetics, linguistics, literature, civilization and TEFL). In this case, English is taught for general purposes to those learners who intend to become specialized in this language in order to get a licence degree (Bachelor of Art). The teaching cycle of the licence is undergone in three years composed of six semesters. The teaching of English at university aims at developing the learners' competencies through the instruction of various subject-matters. The main modules taught to first year and second year EFL undergraduate students are: phonetics, grammar, techniques of oral production, techniques of written production, discourse comprehension, literary studies, linguistics, Anglo-Saxon civilization and African civilization.

The process of evaluation involves two phases. The first phase is human based evaluation resulting from the teachers' assessment which takes the form of marks or grades (a number of points out of 20). The second phase is machine based as the computer is in charge of converting the marks into credits as well as listing those students who obtained the required credits for passing and those who did not. In this way, the procedures of calculating students' grades are achieved via a software system. Regarding the methods of assessing the learners, the examinations frequently rely on the employment of essay questions and the technique of holistic scoring is often involved when assessing students' work.

In fact, the main objective of testing EFL learners is to take a decision concerning their placement in a specific level. Nevertheless, the examination results can lead to a diagnosis of the teaching and learning situation. They can also reveal the degree of the effectiveness of the assessment methods and may be employed as a sort of feedback to help the learners to improve their level.

The students are evaluated for placement or certification purposes. Therefore, the evaluation of first and second year undergraduate students aims at placing these learners in the following levels while the third year students are assessed in order to be awarded a licence degree. The testing process is based on examinations that may involve indirect methods in the units related to the language skills as well as direct methods in the modules related to content areas. In addition to the use of examinations, a form of continuous assessment takes place through the evaluation of the learners' homeworks, projects and training.

### **Case Study**

To evaluate and understand the testing situation, it is important to set up a research design that functions within four dimensions referring to the main elements which interact in the process of assessment. These elements are: the examinees who are the students, the tester who is the teacher, the content of the test or examination and the scoring procedure. To give an idea about the purpose of the present research, one has to list some of the findings concerning the domain of investigation. Among the main problematic questions related to testing, the issue of the content of the examination or the topic effect may be raised. The main research questions that can be asked are as follows:

- How does the topic or examination question affect the learners' degree of achievement?
- What is the relationship between the topic effect, the scoring procedure and the learners' results?

To attempt to answer these questions, the following hypotheses may be put forward:

- Some students may face difficulties in dealing with certain topics
- The teacher's way of scoring may vary depending on the given subject.

The target population involves EFL undergraduate learners. A random sampling method is adopted to ensure an equal probability of sampling. The selected samples include a group of first year students in addition to a group of second year students. Within this study, there are various variables that are examined. The first variable is the students' writing performance which is a dependent variable as it is a changeable element that has to be analysed in relation to other elements. The other variable is the topic effect which represents an independent variable in the sense that it is an element that may lead the other variables to change.

The main research tool involved in this study is document analysis which implies gathering data from the educational materials that are available to the researcher. According to Cohen et al.(2007) document analysis has many advantages as it displays little reactivity from the part of the writer of the document which leads the collected information to be free from bias and distortion. Also, this research tool enables the researcher to study large samples. Thus, the documents used for data collection include the examination topics and students' scores.

This study focused on the case of first year undergraduate students of English with special reference to their results in the module of literary studies. The sample involved seventy-three (73) students. It was necessary to analyse the examination topics and the learners' results in the two yearly examinations. The content of the first term examination of the module of literary studies was as follows:

**Exercise I:**  
**A) Choose the correct answer (2 pts)**  
 1- The Globe referred to .....  
 a) a collection of Shakespeare's plays  
 b) a theatre where Shakespeare worked  
 c) Shakespeare's house  
 2- Figures of speech are used in order to .....  
 a) denote the genre of literature  
 b) provide arguments  
 c) impress and attract the reader  
**B) Find out if these statements are true or false. Then, correct the false ones (5 pts)**  
 1- Fables are stories about everyday life introduced by "once upon a time....."  
 2- A metaphor is an implied comparison that does not use the words 'like' and 'as'  
 3- The ballad is a lyric poem  
 4- The difference between the ode and the sonnet lies in the form of language used by the poet

**Exercise II: Identify the type of the figures of speech included in the following statements (3 pts)**  
 a) All the world's a stage  
 And all the men and women merely players  
 b) Her smile was a mile wide  
 c) Casual dress, like casual speech, tends to be loose, relaxed and colourful.  
 d) The boy is so intelligent that he failed in all the subjects  
 e) Snow had wrapped a white blanket over the city.

**Exercise III: (10 pts)**  
 "Literature manifests itself in various forms to fulfil different functions"  
 Comment (in no more than 12 lines)

Figure 1- First term literary studies examination worksheet

From the above worksheet, it can be inferred that the examination topic is composed of various writing tasks. The first part of the first task is based on the close-ended response format as it provides the learners with multiple-choice answers. The second part of the first activity and the second exercise follow the limited-response format since these tasks require the students to give a limited number of answers. The last exercise takes the form of the open-ended response format since the testees are asked to produce a short essay and are free to express themselves in whatever manner they want. These types of activities impose the use of indirect and direct methods of assessment. The first and second exercises rely on indirect or objective methods of assessment as they represent objective questions that are scored in a mechanical manner through the use of analytic scoring. On the other hand, the third exercise is based on direct methods of assessment or performance assessment since it puts the emphasis on the learners' writing production via the employment of holistic scoring which involves the assessor's opinion about the quality of the given piece of writing.

The scoring scale tends to be analytic in the first and second exercise while it is holistic in the third one. Thus, 50% of the examination questions can be answered and scored in an objective way. On the other hand, the topic of the second examination was as follows.

***Choose one of the following questions:***

*I) Analyze the following passage (your essay should include: identifying the passage and the literary genre, the main ideas, the theme, the setting, two figures of speech).*

The second of June 1608, Smith left the fort to perform his discovery with his company...they crossed the bay to the eastern shore and fell with the isles called Smith's Isles, after our captain's name. The first people we saw were two grim and stout savages, with long poles like javelins, headed with bone. They boldly demanded what we were and what we would, but after many circumstances they seemed very kind and directed us to Accumac, the habitation of their Werowan, where we were kindly entreated. This king was the comeliest, proper, civil savage we encountered. His country is a pleasant fertile clay soil, some small creeks, good harbours for small barks but not for ships. They spoke the language of Powhatan, wherein they made such descriptions of the bay, isles and rivers that often did us exceeding pleasure.

*II) Write an essay in which you speak about the literary forms that characterize colonial American literature.*

Figure2- Second term literary studies examination worksheet

It can be noticed that the topic of the second examination involves a choice between two questions that follow the open-ended response format which yields to direct methods of assessment. The first question is in the form of text analysis requiring the learners to study a passage extracted from John Smith's General History of Virginia, New England and the Summer Isles (1624). Although certain elements of the literary analysis of this text like the genre, the theme, setting and figures of speech require limited or exact responses, the remaining parts of the essay have to be elaborated on the basis of the learners' personal interpretation. Thus, this question has to be answered and scored in a manner that is rather subjective. On the other hand, the second alternative is in the form of a general essay question requesting the students to summarize and synthesize the taught information. Consequently, the rating procedure is holistic. The following table gives an idea about the students' degree of achievement in the two yearly examinations of the module of literary studies.

<i>Grades</i>	<i>2-3</i>	<i>4-5</i>	<i>6-7</i>	<i>8-9</i>	<i>10-11</i>	<i>12-13</i>	<i>14-15</i>
<i>Number of students (1<sup>st</sup> Exam)</i>	01	02	07	28	19	11	05
<i>Number of students (2<sup>nd</sup> Exam)</i>	05	07	12	17	18	11	03

Table1- Students' degree of achievement in the module of literary studies

The above table reveals that thirty-five students, representing 47.94% of the total number, got the average in the first examination while thirty-two (32) learners, representing 43.83% obtained the average in the second examination. Hence, there is not a great difference in the results of those learners whose level is good or average. However, there is a variation in the achievement of those learners whose level is below the average. For instance, 38.35% of the students got grades between 8 and 9 out of 20 in the first examination whereas in the second examination the percentage of the learners obtaining such grades is 23.28%. Moreover, ten (10) students, representing 13.69%, got less than 8 out of 20 in the first examination while twenty –four (24), representing 32.87%, obtained less than 8/20 in the second examination.

Therefore, the nature of the topic can have an effect on the learners' performance. In the second examination, forty (40) students got lower marks than those they obtained in the first examination. Twenty-five (25) learners obtained better grades and eight (8) students kept the same level. The following pie chart explains how first year students' second examination results varied from those of the first one.

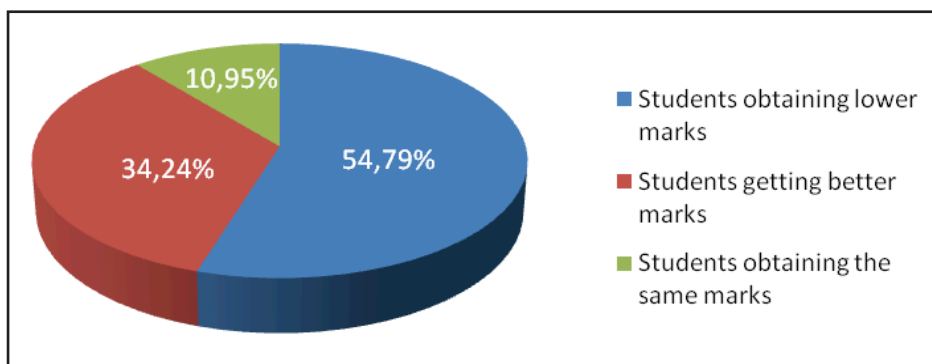


Figure3- The degree of variance of first year students' results

From the above pie chart, it can be inferred that the level of a considerable number of students decreased since the marks they obtained in the second examination were lower than those of the first one. These results can be linked to the factor of topic effect since not all first year undergraduate students are capable of writing good essays in English. Thus, this category of learners may not perform very well in the second examination which is totally based on essay writing. Moreover, the rating scale of the first examination is more objective than that of the second one which relies on the teachers' personal judgment. In this way, the problematic issue in assessing students' essays in the module of literary studies would be related to the design and choice of the examination topic as well as the difficulty of designing an objective evaluation grid that takes into consideration the assessment of the content and the form of the essay.

The issue of the topic effect can be raised in different modules and at various levels. As an illustration, one can mention the case of second year undergraduate students of English. The sample involved the results of a group of twenty-four (24) students in the examinations of the module of Techniques of Written Production (TWP). The first examination included the following questions:



**Exercise 1:**

***Break down the following thesis statements into three sub topics: (3pts)***

- Obesity is becoming a serious problem in some countries
- Reading is a beneficial activity

**Exercise 2:**

***In the following introductory paragraph, the sentences are in an incorrect order.***

a) Rewrite the paragraph following the requirement of an introduction (4pts)

b) Write the whole essay (13pts)

- They demand that the air they breathe, the water they drink and the food they eat is free of pollution and contaminants.
- In recent decades, there has been a growing realization that the quality of the air, water, soil and food affect our health and our lives.
- A clean and healthy environment is part and parcel of the wealth and quality of life that we desire for ourselves now and our children in the future.
- This is why people want to live in a healthy environment

Figure 4- First term TWP examination worksheet

Therefore, the questions of the first examination do not confine the students to the production of an essay and they can be scored through a combination of an analytic and holistic marking scale. On the other hand, the topic of the second examination involved an essay question which was formulated as follows:

**Choose one of the following topics:**

- 1- Explore and analyse the effects of the cell phone on society and interpersonal relationships.
- 2- Discuss the following statement:  
*Any student caught cheating on an examination should be automatically dismissed from college.*

Figure 5- Second term TPE examination worksheet

The above topic relies exclusively on essay writing. It requires a holistic procedure of scoring and involves a larger degree of subjectivity from the testees and the scorer. The results of the learners in the two examinations differed to a certain extent. In the first examination,

fifteen (15) students, representing 62.50% of the total number, obtained the average; in the second examination, ten (10) students, representing 41.66% , got the average.

The following pie chart summarizes the variance in students' second examination results from those of the first one.

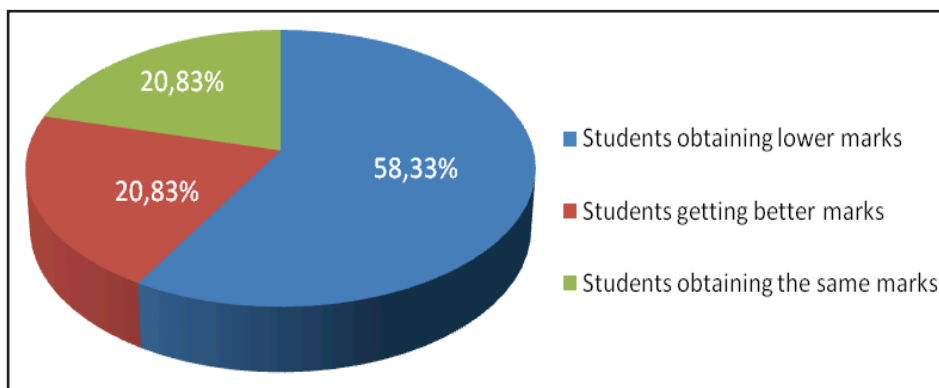


Figure 6- The degree of variance of second year students' examination results

From the above table, it can be stated that the students' level decreased to a certain extent in the second examination since fourteen (14) students obtained lower marks than those they got in the first examination; five (5) learners obtained better grades and five (5) students got the same marks as the first examination.

## Discussion

Generally speaking, the nature of the topic can affect EFL learners' examination results for two main reasons. The first one is related to the learners' degree of appreciation of the topic, their knowledge of the topic and the level of their writing performance. Hoffman (2003) mentions the problem of relevance since "...thinking of suitable topics is not easy. Such people are discriminated against if the list of topics contains nothing that stimulates their creativity" (p. 44). The second reason refers to the fact that each topic requires a specific rating scale that urges the teacher to score the answers either objectively or subjectively.

It can be asserted that the topic or examination question may affect the learners' degree of performance. Hence, one would observe a high or low level of achievement depending on the nature of the questions involved in the examination. Nevertheless, the topic effect is not the only factor that determines students' success or failure in answering the examination question.

According to Wainer (2007), "observed score = true score + error" (p.15). This means that the true score is the score that would be obtained by the learner if the test is administered several times while the error score refers to the degree of the test's unreliability. As a result, the error refers to the factors that may cause the results of the examination to be unreliable. This kind of error in results may be linked to the learners' abilities, their mood and health as well as the content of the examination, the test taking conditions and the rating scale.

In fact, various elements interact in the process of assessment including the testees, the scorer and the scoring procedure. The testees may achieve a high or low performance depending on their linguistic background and the degree of their appreciation of the topic. Also, there are certain topics that have to be answered subjectively. For instance, "Literary works are so complex that there is no single right answer about their meaning" (Acheson, 2010, p. 83). This means that certain examination questions are subjectively answered and scored.

Moreover, a set of cognitive, psychological and social factors may affect the learners' achievement. For instance, a student may face problems in learning and performing well because of certain factors like the degree of concentration and attentiveness which can influence the function of the memory and may cause a problem in remembering information (Latash, 2008, p.180). Besides, the learners' degree of motivation, interest and mood can have an impact on their achievement. In addition to this, the stimulus and encouragement provided by the parents, the teachers and the educational institution may determine the learners' success or failure.

On the other hand, the scorer's subjectivity can determine the quality of the examination results. For instance, the scorer may stick to a specific model of answering the essay question and assume that all the learners should provide the same version. In certain cases, students' handwriting can influence the rater's judgement. Moreover, there is a close relationship between the scoring procedure and the learners' examination results since essay questions involve holistic scoring that relies on the scorer's personal judgement which may yield to a certain extent of subjectivity.

## Suggestions

In order to improve the testing situation it is necessary to take into consideration the following suggestions.

The choice of the topic is very important as it may limit the type of language used by the testees (Lynch, 2003). Therefore, the teacher has to select carefully the question of the examination since the latter can affect the learners' performance. As a result, the specification of the content and the evaluation of the examination as well as the required response length can help the learners to construct a point of view about the expected response. Moreover, the teacher may give the learners the possibility to select a question to be answered from a set of alternatives since "having choice is motivating" (Pollack Day et al, 2002, p.24).

Also, the teacher can construct essay questions that assess specific cognitive skills according to Benjamin Bloom's Taxonomy of Educational Objectives. The table below gives an example of a set of essay questions that can be used to test various types of cognitive abilities.

<i>Bloom's cognitive levels</i>	<i>Examples of essay questions</i>	<i>Assessment objectives</i>
Knowledge	- Identify the different teaching methods.	Testing the students' ability to remember concepts and define them
Comprehension	-Explain the following statement: "Teaching is an art"	Assessing the students' ability to explain and summarize what has been taught to them.
Application	-Use the different language properties to explain the difference between human and animal communication	Testing the learners' ability to make use of the taught material to clarify a certain concept
Analysis	- Compare the style of Jane Austen and Charles Dickens	Assessing the students' ability to make a distinction between various concepts by providing arguments.
Synthesis	- Write a short story using the ideas provided in the given poem.	Testing the learners' capacity to develop ideas and produce creative pieces of work
Evaluation	- Evaluate the degree of effectiveness of the competency based approach.	Assessing the students' ability to criticize certain ideas on the basis of the taught material

Table 2- Examples of essay questions

Since holistic scoring relies on the teachers' personal judgement, training novice scorers to score essays consistently is very necessary. Such training enables them to acquire more experience and increase the reliability of scoring (Colburn, 2003). Rater training may take place through the moderation meeting which implies requiring novice and experienced raters to score papers, making a discussion regarding the distinction between the scores of the trained and experienced scores, giving instructions to the novice raters to moderate their way of scoring and asking them to rate real essays (Lynch, 2003).

In addition to this, teachers may decrease subjectivity by utilizing analytic scoring which implies giving separate scores to the different components of the learner's essay. Luongo-Orlando (2003) suggests the use of an analytic rating scale based on six writing criteria including the writing process, content, organization and development, visual presentation and format, language usage and application of language conventions. The following writing evaluation sheet can be employed to rate essays in an analytic manner.

<i>Criteria</i>	<i>Rating scale</i>				
	5 points	4 points	3 points	2 points	1 point
Content					
Organization					
Word choice					
Language conventions					

Table 3- Writing Evaluation Sheet

In fact, the above evaluation writing sheet relies on the assessment of four writing criteria involving the content of the essay, organization, word-choice and language conventions. The score of each criterion is five points. The criterion of content includes the provision of clear ideas, accurate information, explanations and adequate details. The feature of organization means a structured essay following logical development. The criterion of word choice implies the use of rich vocabulary and effective word choice in the essay. The scoring of language conventions puts emphasis on the testee's knowledge of spelling, mastery of grammar as well as the use of capitalization and punctuation.

In fact, various types of writing performance should be evaluated when scoring essays (Leo and Murphy, 1988). Moreover, feedback is necessary since it allows the learners to know their strengths and weaknesses. In this way, the testees recognise the types of mistakes they produced in the examination, analyse the reasons behind such deficiencies and think about the ways that will lead them to improve their level. Kozma et al. (1978) suggest the use of a profile sheet as a feedback activity. As an illustration, the following profile sheet (adapted from Kozma et al., 1978, p. 103) can be utilized to inform the learners about the test's objectives and their level of achievement.

Objective	Not attempted	Incomplete	Demonstrated mastery
Introducing the topic			*
Defining concepts		*	
Use of appropriate vocabulary		*	
Provision of explanations and arguments		*	
Use of examples and illustrations	*		
Use of quotes	*		
Providing a conclusion			*
Respecting language conventions			*

Table 4- An example of a profile sheet

Therefore, a profile sheet may enable the learners to learn from their mistakes. In addition to this, the teacher has to make research about the relationship between the teaching and testing situation. The learner's scores have to be analysed to identify the sources of error in the examination results and to increase the test's reliability. The degree of error in the results may be reduced by providing good testing conditions, designing appropriate examination questions and decreasing the scorer's inconsistency.

## **Conclusion**

The present article attempted to provide an idea about the different elements that interact in the process of assessment as well as the factors that may enhance or hinder the learner's writing performance. The research results revealed the existence of different factors that influence the learners' results including the topic effect. As a consequence, a set of suggestions and recommendations were provided. These suggestions focused on the specification of the criteria that should be involved in the selection and the construction of the examination questions and topics in addition to the design of scoring rubrics and the elicitation of the principles that should be taken into consideration when training the scorers to rate students' essays.

Finally, it has to be stated that the testing process remains complex because it is based on the interaction between different variables that cannot be totally controlled. Thus, further research is needed concerning the assessment of EFL learners.

## REFERENCES

- Acheson, Katherine. (2010). *Writing Essays about Literature: A Brief Guide for University and College Students*. Canada: Broadview Press.
- Bloom, Benjamin. (1956). *Taxonomy of Education Objectives*. New York: Longman.
- Cohen, Louis, Manion, Lawrence, & Morriison, Keith. (2007). *Research Methods in Education*. London: Rutledge.
- Colburn, Alan. (2003). *The Lingo of Learning: 88 Education Terms Every Science Teacher Should Know*. USA: National Science Teachers Association.
- Hoffman, Banesh. (2003). *The Tyranny of Testing*. New York: Dover Publications.
- Kozma, Robert, Belle, Lawrence, & Williams, George Warner. (1978). *Instructional Techniques in Higher Education*. New Jersey: Educational Technology Publications.
- Latash, Mark. (2008). *Neurophysiological Basis of Movement*. USA: Human Kinetics.
- Leo, Ruth, & Murphy, Sandra. (1988). *Designing Writing Tasks for the Assessment of Writing*. New Jersey: Greenwood Publishing Group.
- Luongo-Orlando, Katherine. (2003). *Authentic Assessment: Designing Performance-Based Tasks*. Canada: Pembroke Publishers.
- Lynch, Brian. (2003). *Language Assessment and Programme Evaluation*. UK: Edinburgh University Press.
- Pollack Day, Jeni, Spiegel, Dixie Lee, Mclellan, Janet, & Brown, Valerie. (2002). *Moving Forward with Literature Circles: How to Plan, Manage and Evaluate Literature Circles that Deepen Understanding and Foster a Love of Reading*. New York: Scholastic Inc.
- Wainer, Howard, Bradlow, Eric, & Wang, Xiaohui. (2007). *Testlet Response Theory and its Applications*. Cambridge: CUP.