

Evaluating Examination-Papers of EFL Students at Hodeidah University-Yemen

By

Dr. Mariam Youssef Mohammed Abdu

د. مريم يوسف محمد عبده عبد

Abstract

This study mainly throws light on the quality of examinations undertaken by English students in the English Department, Faculty of Education at Hodeidah University. It aims at evaluating the types of the questions included in the examination papers in order to find out whether the examiners put into account the different stages of complexity proposed by Bloom (1956). Eighty examination papers have been collected randomly and examined with respect to Bloom's taxonomy. It is found that most questions in the examination-papers focus on assessing lower thinking skills such as knowledge, and comprehension category. Questions which involve application of knowledge are, also, significantly, high. However, questions assessing higher abilities, i.e. synthesis and evaluation are not assessed at all.

Keywords

Bloom's taxonomy, examination, achievement tests, lowers-order thinking skills, higher-order thinking skills

1. Introduction

1.1 Importance of Tests in Language Teaching

Tests and examinations are tools to measure the level of knowledge, understanding and ability to apply knowledge. They are very important tools for educators to determine the quality of the teaching- learning process. According to Richards and Schmidt (2002: 546), a test is defined as “*any procedure for measuring ability, knowledge, or performance*”.

Teaching and testing a language are so much interrelated. In other words, the quality of the test determines the quality of teaching materials and teaching process. As Alderson and Waters (1983:43) have observed that “*. The test will exert an influence on the teaching regardless of whether it is a bad test or a good test. So, if a bad test can have a bad effect on the teaching, a good test can have a good influence on teaching. And since the test will, always, in our experience, have some effect on teaching, the simplest way to innovate language teaching is to improve the test.*” Tests are the instruments of evaluation which measure the students' performance. Hence, appropriate questions should be selected to assess the students' performance through the required complexity levels.

Assessments, in the form of written tests, are necessarily used in the educational process. Tests are significant in the sense that they provide information about the achievement of students. In addition, tests provide information about the extent to which course outcomes are achieved. Without such information, as Hughes (1989:4) points out, “*it is difficult to see how rational educational decisions can be made*”.

Tests are classified according to the types of information they provide and to different purposes. Although there are different types of tests i.e. achievement tests, proficiency tests, placement tests and diagnostic tests, the present study exclusively focuses on achievement tests. Hughes (1989:13) defines achievement tests as those tests which are “*directly related to language courses*”. He identifies the purpose of such tests as establishing the extent to which individual students or groups of students have achieved the course objectives. The main objective of carrying out examinations in the EFL context is to assess the knowledge and the skills taught at a specific period of time. Heaton (1975:5) claims that “*Tests may be constructed primarily as devices to reinforce learning and to motivate the students or primarily as a means of assessing the student's performance in the language*”.

The terms test and examination overlap and they are, frequently, used together as a lexical 'chunk'. Ahmad (2001:2-6) illustrates a test as "*a set of questions, problems or exercises to which the students are asked to respond to obtain an appraisal of designated characteristics of the students such as specific kind of knowledge, aptitude, abilities and skills; in education used interchangeably with examination*". Examination is mostly equated with testing because the function of both systems is to evaluate and assess the students' performance. In this study, the term examination is used to refer to the achievement tests which are used to measure the students' knowledge and performance at the end of the courses in universities.

1.2. Importance of the Study

Evaluation is an unavoidable procedure in any teaching- learning scenario. By evaluating a process or a programme, we intend to check their effectiveness and whether their purposes have been achieved. In this context, Hutchinson & Waters (1987:110), define evaluation as "*a matter of judging the fitness of something for a particular purpose*". Further, Gronlund(1981:5) defines evaluation as "a systematic process of determining the extent to which instructional objectives are achieved by pupils." Thus, evaluation is an important part of the educational process. It plays an essential role in education because it can bring useful and helpful information about the quality of the teaching –learning process. Given this fact, it is apparent that carrying out an evaluative study of examinations undertaken by EFL learners is necessary to identify their qualities and to highlight the weaknesses in designing them.

2. Problem of the Study

At the end of every examination period, many students of the English Department in the Faculty of Education at Hodeidah University complain that the questions in the examination-papers are inadequate to evaluate their levels. Some students claim that the questions are very difficult and beyond their levels. Others say that the questions are very easy and less than what they expected. Many students, also,

claim that some examination-papers focus entirely on one type of questions which requires rote memorization such as recalling information about dates, events, places and major ideas through testing items like multiple-choice, defining, and describing. Thus, only those who can memorize a large amount of information pass the examinations successfully. In this sense, many students think that examinations don't evaluate their actual levels.

In general, examinations are very common assessment tools used in universities. Since written examinations, in particular, are the major and sometimes the only source used for assessing the students' performance in the EFL context in the Yemeni universities, it is of a great importance to study the types of the questions included in examinations in order to pinpoint the source of such problem.

3. Purpose of the Study

The present study attempts to shed light on the quality of questions included in the examination-papers given to students of English in the English Department at Hodeidah University. It aims at evaluating the types of the questions included in the examination-papers in order to find out whether the examiners put into account the different stages of complexity proposed by Bloom (1956) while designing the examination-papers. Such an evaluation, as hoped, may help in improving the quality of the testing items included in examinations given to students of English and, consequently, improving the teaching and learning processes.

4. Questions of the Study

The present study tries to answer the following questions:

1. To what extent is the availability of Bloom's higher thinking levels in the examinations held in the English Department at Hodeidah University?
2. Do the teachers of the English Department include Bloom's cognitive goals levels in the examinations held in the Department?

5. Research Methodology and Tools

This study is mainly based on Bloom's Taxonomy for complexity levels. It has employed the document analysis research method. The sample of examination papers has been collected randomly. Eighty examination-papers have been collected randomly. The examinations have been undertaken by students in the English Department at Hodeidah University along five successive years starting from 2009 until 2014. They have been analyzed according to Bloom's Taxonomy to determine the levels of complexity. The results then have been tabulated and the frequency of each complexity level is recorded and discussed in relation with Bloom's Taxonomy of cognitive domain.

6. Population:

Available literature in shape of books, research articles, research theses, question papers are the materials of the study. Samples for the study are taken from question-papers given to students in the Department of English at Hodeidah University. The study is mainly conducted with reference to Bloom's taxonomy of educational objectives.

7. Limitation of the Study

This research is confined to examination papers carried out in the English Department in the Faculty of Education at Hodeidah University in five successive years beginning from 2009 until 2014. Since the evaluation is directed towards an EFL context, the study is solely restricted to the courses of English major. Requirement subjects that are taught in Arabic such as Islamic studies and Arabic language are excluded.

8. Sample of the Study

The sample of the study has included the examination-papers given to students in the English Department in four levels along five successive years beginning from 2009 until 2014. Three hundred and sixty two examination-papers have been collected. Twenty examination-papers have been selected randomly from each level and a total of eighty examination papers have finally been selected.

9. Why Using Bloom's Taxonomy

Bloom's Taxonomy has been applied to a variety of situations. It is proved to be helpful and widely accepted in almost all circumstances to describe objectives and procedures of a learning process. Bloom specified six levels of complexity starting from the lowest to the highest: knowledge, comprehension, application, analysis, synthesis, and evaluation. It includes three overlapping domains: the cognitive (intellectual), psychomotor (physical), and affective (attitudinal)



Figure (1): Bloom's Taxonomy of Educational Objectives

This classification is used as the standard for judging whether tested items are Lower-Order-Thinking (LOT) which include knowledge and comprehension, or Higher-Order-Thinking (HOT) which include analysis, synthesis and evaluation. Application often falls into both categories.

Verb and question examples that represent intellectual activity on each level are listed below:

1. **Knowledge:** arrange, define, duplicate, label, list, memorize, name, order, recognize, relate, recall, repeat, and reproduce state.

Question verbs: Define, list, state, identify, label, name, who? when? where? what?

2. **Comprehension:** classify, describe, discuss, explain, express, identify, indicate, locate, recognize, report, restate, review, select, and translate.

Question verbs: Explain, predict, interpret, infer, summarize, convert, translate, give example, account for, paraphrase x?

3. **Application:** apply, choose, demonstrate, dramatize, employ, illustrate, interpret, operate, practice, schedule, sketch, solve, use, write.

Question verbs: How could x be used to y? How would you show, make use of, modify, demonstrate, solve, or apply x to conditions y?

4. **Analysis:** analyze, appraise, calculate, categorize, compare, contrast, criticize, differentiate, discriminate, distinguish, examine, experiment, question, and test.

Question verbs: Differentiate, compare / contrast, distinguish x from y, how does x affect or relate to y? why? how? What piece of Y is missing / needed?

5. **Synthesis:** arrange, assemble, collect, compose, construct, create, design, develop, formulate, manage, organize, plan, prepare, propose, set up, and write.

Question verbs: Design, construct, develop, formulate, imagine, create, change, write a short story and label the following elements:

6. **Evaluation:** appraise, argue, assess, attach, choose, compare, defend estimate, judge, predict, rate, core, select, support, value, evaluate.

Question verbs: Justify, appraise, evaluate, judge according to given criteria.

During the late 1990's, Anderson and Sosniak provided a revised version of Bloom's Taxonomy. The subcategories of the six major categories were replaced by verbs, and several subcategories were reorganized. Their taxonomy, also, included six levels. These are: remembering, understanding, applying, analyzing, evaluating and creating.



Figure (2): Design of Anderson's Taxonomy

Wilson's classification of levels in the cognitive domain is considered as an extension of Bloom's taxonomy. In Wilson's model, the objectives in the cognitive domain are classified into four levels: remembering, understanding, application, and analysis.

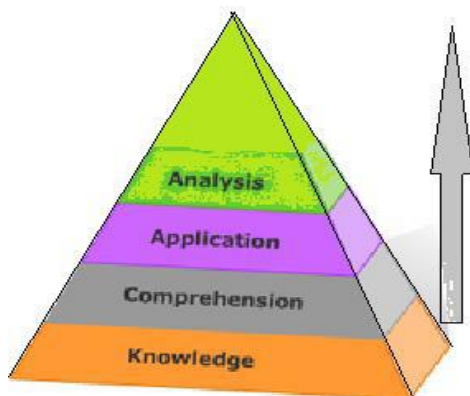


Figure (3): Design of Wilson's Taxonomy

Nayef, Ismail and Yaacob (2013:174), who carried out a comparison among these three taxonomies, claim that Bloom's taxonomy is more suitable as an analysis tool to Educational Objective domain than the remaining revised versions of Bloom's Taxonomy. They state that "Bloom's Taxonomy has proved its superiority and effectiveness over other methods and it could serve as a common language about learning goals"

10. Review of Related studies

There has been a considerable amount of research on the evaluation of examination papers based on Bloom's Taxonomy. Reviews of research in the United States, the United Kingdom and in many developing countries have shown that questions asked by teachers in class or in tests are of a low intellectual level, usually requiring little more than rote memory and recall.

Beginning with Bloom (1956) who carried out analysis on classroom questions and found that over 95% of the test questions that students encounter require them to think only at the lowest possible level; that is knowledge which requires the simple recall of information.

Mohnot (2006) carried out a study aimed at examining the use of Bloom's Taxonomy of Educational Objectives (Cognitive Domain) in

the Indian Certificate of Secondary Education (ICSE). This study categorized the skills required to solve ICSE English Literature exam questions at four points (phases) during the lifetime of the exam. He found an increase in the questions of low order of thinking level and a decrease in the high level ones. Also, 90% of the questions were at two lowest levels (knowledge and comprehension).

Recent studies in evaluating examinations show the same results. Swart (2010) carried out an evaluative study on the examination papers of the Engineering section and found out that these examinations focused on low order thinking questions. He concluded saying that, "*Engineering academics in this case study may be relying too heavily on LOq when they should be using more HOq in their final examination papers...*" Alzu'bi(2011) carried out an analytical study of the Community College Associate Degree Examination based on Bloom's cognitive levels. He discovered an increase in knowledge-based or lower order thinking questions and

no substantial rise in higher order thinking questions. His analysis showed that nearly 90% of the questions in all phases fell into Bloom's two lowest categories, knowledge and comprehension. Sufiana (2012) conducted an analytical study of question papers of Pakistan Studies for secondary level with reference to Bloom's taxonomy of educational objectives. She depicted that maximum focus in examinations was on knowledge, and comprehension category and whole findings also reflected that application, analysis, synthesis and evaluation were not assessed.

The researcher of the present study has tried to shed light on the quality of questions included in the examination papers in order to see whether the findings of the present study would be consistent with the findings of the previous evaluative studies of examination papers

which were, also, based on Bloom's Taxonomy of educational objectives.

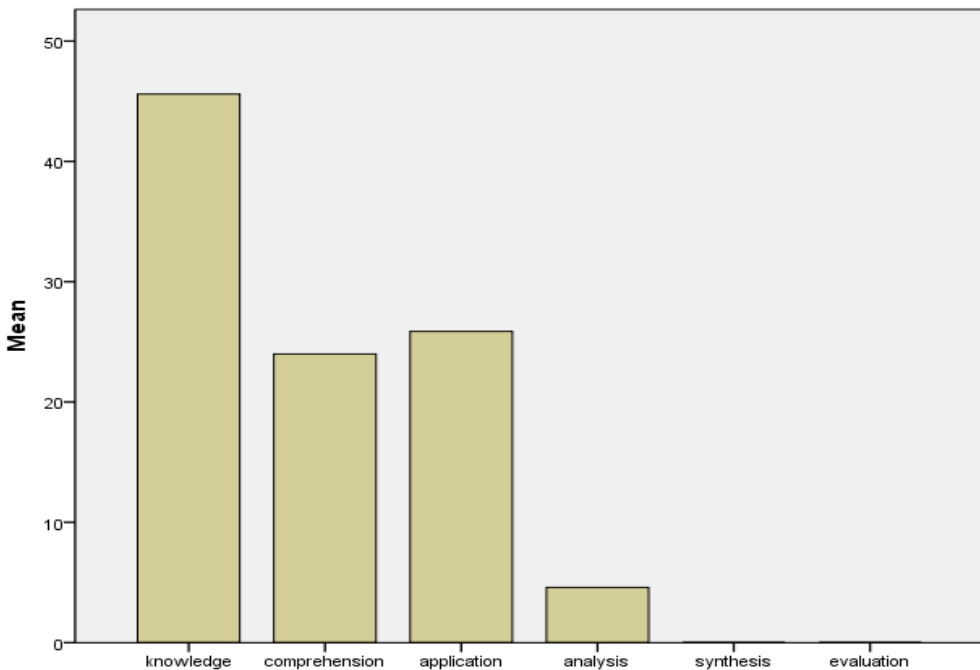


Figure (4): Graphic representation of the distribution of the thinking-levels in examinations

11. Discussion

According to the analysis of the data, it is clear that instructors tend to avoid asking questions which require high- order thinking level. Instead, they badly focus on the low-order thinking levels.

When comparing the first two levels which involve lower-order thinking skills, i.e. Knowledge and comprehension to the remaining levels which involve higher cognitive skills, it was apparent that knowledge, and comprehension questions were significantly higher than those for analysis, synthesis and evaluation: The mean for the knowledge level was ($M=45.599250$), The mean for the Comprehension level was ($M=23.989875$). However, questions which involve application of knowledge were, also, significantly high compared to the remaining higher levels ($M=25.8734$). Though, the mean for the analysis level was ($M= 4.5966$), it was rather very low compared to the previous levels. But this indicates that, at least, there were some questions related to this high level of cognitive domain.

Also, the study reflected that questions which require higher cognitive skills such as synthesis and evaluation were, surprisingly, unavailable. There was no single question which tests these skills among the sample selected. This means that the results of this study support the findings reached by the researchers of aforementioned previous studies.

12. Conclusion and Recommendations

Thus, the findings of the present study are consistent with those of the previous studies focused on evaluation of examination-papers, depending on Bloom's Taxonomy of educational objectives.

In the light of the reached findings, it is clear that instructors at Hodeidah University tend to avoid asking questions which require high- order thinking level and badly focus on the extremely low-order thinking levels. This makes it urgent for the stakeholders in the Ministry of Higher Education to rethink of the way examinations are designed and find suitable solutions for the poor quality of

University examinations in order to achieve the education goals prescribed by each university.

The following recommendations are suggested by the researcher in order to improve the quality of examination papers in the English Department at Hodeidah University:

- 1) Instructors should put into account Bloom's taxonomy of Educational Objectives while preparing questions because higher - order examination questions lead students to think creatively. Examinations should include questions targeting higher order thinking more.
- 2) Well-designed testing items incorporating different learning levels of Bloom's taxonomy may be a potential method of assessing critical-thinking skills. Well- prepared and reasonable examination papers should include various complexity levels to accommodate the different capacities of students.
- 3) There should be agreement between how instructors teach and how they assess the level of questions they ask.
- 4) A centre for quality and improvement of the teaching –learning process may be established in the university in order to evaluate the teaching and assessment process. All instructors should be required to submit course reports on the course(s) they teach. Such reports should include the course objectives based on Bloom's taxonomy, contents of the course and assessment methods. The examination questions should be, also, related to the course objectives.
- 5) There should be effective programs and workshops for English staff members to improve their knowledge and skills in asking higher order-thinking questions.

11. References

- 1-Ahmad, M. (2007). *Principles of Education*. IlmiKitabKhana, Lahore, Pakistan.
- 2- Alzu'bi, M. (2011). Analytical and Evaluative Study of English Questions in the Community College Associate Degree Examination in Albalqa Applied University from 2007 to 2010. *Proceedings of the 3rd International Conference of Teaching and Learning (ICTL 2011)* INTI International University, Malaysia.
- 3- Anderson , L.W., & Sosniak, L.A. (Eds.). (1994). *Bloom's taxonomy: a forty-year retrospective. Ninety-third yearbook of the National Society for the Study of Education*, Pt.2 . , Chicago , IL . , University of Chicago Press.
- 4- Bloom, B. et al.(1956). *Taxonomy of Educational Objectives: The Classification of Educational Goals: Handbook I Cognitive Domain*. David McKay. New York.
- 5-Edwards J.(1991). *Evaluation in Adult and Further Education: A Practical Handbook for Teachers and Organizers*. Liverpool. UK. The Workers Educational Association
- 6-Heaton,J.B.(1975).*Writing English Language Tests*. London: Longman.
- 7-Gronlund, N. 1981.*Measurement and Evaluation in Teaching*. New York: Macmillan.
- 8-Hutchinson and Waters. (1987). *English for Specific Purposes: A learning-centered Approach*. Cambridge University Press.
- 9-Hughes, A.(1989).*Testing for Language Teachers 2nd Edition*. Cambridge: Cambridge University Press.
- 10-Malik.(2012). An Analysis of Pakistan Studies Question Papers at Secondary Level *Interdisciplinary Journal of Contemporary Research in Business*. VOL 4, NO 5.
- 11-Mohnot, A. (2006). 'Seeing the Bigger Picture': *Higher Order Cognition in the Indian Certificate of Secondary Education (ICSE) English Literature Examination*. (Master Degree). Stanford University.

- 12-Nayef, E., Yaacob, N and Ismail, H.(2013). Taxonomies of Educational Objective Domain. *International Journal of Academic Research in Business and Social Sciences*. Vol. 3, No. 9 ISSN: 2222-6990
- 14-Richards, Jack C. and Richard Schmidt.(2002). *Longman Dictionary of Language Teaching and Applied Linguistics*. London. Pearson Education Limited.
- 15-Swart, A, J.(2010). Evaluation of Final Examination Papers in Engineering: A Case Study