

TYPES OF WRITTEN EVALUATION FOR ESP AND SCORE ANALYSIS (A CASE STUDY)

Alina BALAGIU¹

Marioara PATESAN²

¹Associate Professor PhD, Electrical Engineering and Naval Electronics Department, “Mircea cel Batran” Naval Academy, no 1, Fulgerului street, Constanta, alinabalagiu@yahoo.com

²Professor PhD, Department of Public administration, economic and social sciences, “Nicolae Balcescu” Land Forces Academy, 3-5, Revolutiei Street, Sibiu, Romania

Abstract: *The paper is a case study based on the hypothesis that the students' scores at written evaluations during a semester should be the same, as long as they are based on the same knowledge the students from different groups gathered. The number of English classes is the same and the subject has been taught in the same manner for all the groups. The analysis is done on groups in the first and second year of study.*

Keywords: *evaluation, ESP, methodology*

Introduction

Teaching English for engineering students implies, as everybody knows, mainly vocabulary and besides some grammar and structures. For learning vocabulary, the ESP student has a more difficult task than a native one because he has to understand the concept and at the same time label it. Hence the principles of teaching and learning vocabulary are frequent exposure to technical words, and not only, and repetition in order the students to memorize the specific vocabulary or terminology. In order to check learning or memorizing several tests should be rendered each term and these types of tests are called progress tests or achievement tests and they should reflect the nature and content of the course. That is way the best English tests for different specialties are those written by the teachers themselves, taking into account their ability in making the tests and the availability of other specialty tests. In designing a good test, there are some basic principles that should be followed:

- a) “Test what you can reasonably assume the learners have learnt”.(1)
- b) Decide what you need to be tested or what the students need to be tested for.
- c) Try to make “objective items” that “can be marked very quickly and completely reliably”. (2)

For writing tests, there are several types of exercises or items suitable for ESP and at the same time in accordance with the principles mentioned above. Purpura (2010) discusses the test tasks according to the type of response that can be “selected response, limited-production and extended-production tasks”. (3) He presents a list of some of the more common testing activities

categorized according to type of expected response. In the selected-response task category are included: “multiple-choice activities, true/false activities, matching activities, discriminating activities, lexical list activities and noticing activities”. (4) In our opinion it is the category with the most objective items or activities that can be rendered to engineering students. From the other two categories only information-transfer and information-gap activities might be used for designing a test for our students.

The first one is the **multiple-choice** completion that may be used both for vocabulary and grammar and that has the advantages of “providing natural context”, “easy and consistent scoring” and “measuring achievement sensitively” (5) and some limitations consisting in preparing the test and students cheating.

Another type of test is **short-answer questions** with the advantages that for the teacher it is quick and easy to grade and for the students it is also quick and easy to write, and we cannot see any disadvantages.

Matching questions are used in tests when you need to assess the knowledge gained from a course. As a rule, with matching test questions, learners get partial credit for answers that are only partially correct. Among the advantages the most important is the fact that a matching exercise gives an objective assessment of the learners' knowledge. Another advantage is the little time used by the teacher to make the test. It can be disadvantageous if there are too many items to be matched and the student might be confused.

Open book tests are those tests where students can use the books or copybooks they have. The greatest advantage is that they are used for assessing the connections a student can make

not his memorizing skills. Among the disadvantages we can mention the fact that some students might simply copy from the book if the tasks are too simple. In Romania such open book tests are not very much used. In our opinion a test of this type once a year might be useful for students who do not have a very good memory.

Tests analyses

The actual paper is a case study based on the hypothesis that the students' scores at the written evaluations during a term should be the same, no matter what type of test paper they are rendered to as long as some conditions are fulfilled. In our opinion the conditions should be:

- the same number of classes for each group or subgroup of students;
- the same information rendered by the teacher which means the same amount of knowledge gathered by the students;
- the same course-book the students can study;
- the same tasks and exercises done in the class or at home.

For the first year of study, specialty engineering, we used three different tests for the vocabulary and about the same type of exercise for the grammar. The tests were rendered at three groups of about 15 students, for each group a different test.

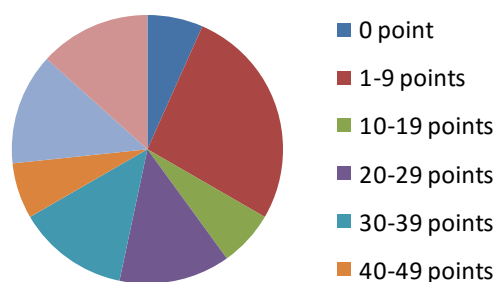
Test one

The first test is mainly a theoretical one being formed of 18 questions from two units studied (Shipping and Ship Types). The expected responses referred to classifications of ships, definitions, characteristics of certain ships, methods of loading and the meaning of some abbreviations. According to the difficulty and the complexity of the question, the score for each answer is between 1 and 5, so that the vocabulary has 70% of the whole test while 30% are given for the grammar activities. Taking into account the difficulty of a theoretical vocabulary test for the students in the first year and with English as a second language, we decided for an open book test. Students could use their materials to answer the questions: course books and notes taken at the seminars, but no their mobile phones. Examples of the questions are:

- What are heavy-cargo ships?
- What does the abbreviation VLCC stand for?

Open book tests are supposed to have better scores than other tests as long as the students have the source of inspiration. The test was taken by 15 students of different English levels starting from beginners up to advanced or starting from A1

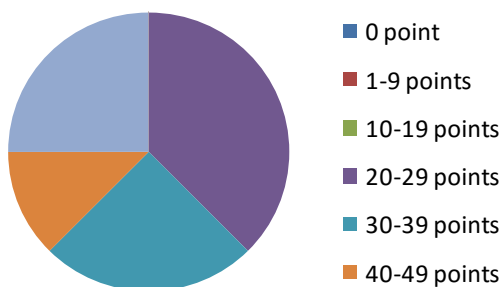
to C2. In order to give a fair mark and to be easier for the teacher to score the test, it has a total of 100 points that is at the end divided by 10. The scores ranged from 2 to 9, however, there was no point given out of 10. At this point we are interested by the 70 points for the vocabulary, so we are going to leave the grammar out. One student did not do anything at vocabulary while the best score was 68. A chart shows the results according to the points taken at the vocabulary for the answers.



The greatest part of the pie chart is represented by people scored up to 10 points at vocabulary, however, also 4 students scored from 50 to 70 points.

Test two

The second test is a combination between 2 theoretical questions, a matching exercise, one task for recognizing abbreviations and a lexical list to be translated. The vocabulary part of the text is also 70% and the grammar is 30%, the same percentage as for the first test rendered. This time the students could not refer to any materials, but their own knowledge. The test was rendered to a number of 16 students. The overall scores were between 2 and 7; the same criterion with 100 points divided at the end by 10 to give a mark to correspond the Romanian system was used. The matching exercise was quite difficult because the students had to match a type of ship with a characteristic, for example: **heavy-lift ships** have as a characteristic **boom** term that can be found in the second column, or **tankers** should have been matched with **carries crude oil and edible oil**. One should have known the traits of the ships very well to solve the exercise, or should have been done the exercise thoroughly during the class. The test is not very easy, but all the tasks and exercises had been done or reviewed before, so that it was a matter of logical thinking and remembering. If we are to demonstrate the score at vocabulary using the same type of chart as for the first test, it would appear as follows.



We can see a great difference from the first chart and this is due to the fact that there are 3 sectors where no student had scored, so that the first scores for vocabulary are in the interval 20-29 points for a number of 6 students. Not less than 4 students scored between 50 and 59, while in the last group there was no student scoring over 60.

Test three

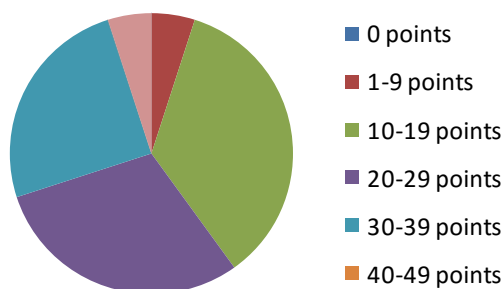
The last type of test is perhaps the most difficult of all as long as our students claimed that the multiple choice tests with multiple answers were considered very “tricky”. The test contains 14 multiple choice items, an exercise with two tasks: match the pictures of ships with their names, and give three characteristics for each type of ship; the last exercise is a lexical list of 16 words connected to ships and navigation to be translated into Romanian. The vocabulary exercises sum 70 points while the grammar part has 30 points, so that the percentages to be the same for a balanced comparison of the results. An example of the multiple-choice item is:

- Any large floating vessel capable of crossing open waters is called

The answers given at this item are:

- a boat; b. a ship; c. a raft; d. a vessel.

The correct answers are in this situation b. a ship and the synonym of ship which is the answer d. a vessel. For the matching exercise, some of the characteristics could have been given just by looking carefully at the pictures of the ships and depicting what is different for eachone. There were 20 students sitting the test, and the results in the form of the chart are as follows.



The third pie chart is different from the other two, the biggest part is represented by the students achieving from ten to 39 points, while at the extremes there is one student with the lowest score and one with the best score. That means from a total of 20 students 18 scored medium and lower.

CONCLUSIONS

We started the paper from the hypothesis that students' scores at written papers should be the same as long as the same conditions are fulfilled. The tests rendered were different and the results were visible different. The easiest test in our opinion was the first one, the open book test, where we had expected the best results. The chart shows us that the lowest score was registered at this test, a student who did not do anything at vocabulary, although he could have used the information (perhaps he has no available source). The largest portion is represented by students that scored between 1 and 9 points at vocabulary, with a total of 4 students. On the other hand the chart is balanced, with a wide range of scores from 30% to 100%.

The next two tests have a predominant large area, the second one at 20-29 points, meaning about 30% of what should have been done, representing the knowledge of 6 students, and the third one with 7 students scoring between 10 and 19 points, meaning about 15% of the total. Taking into consideration the fact that there is only 1 student with 100% we can draw the conclusion that either the last test was the most difficult or the group given the test prepared the least for the test.

- (1) Hutchinson, T., Waters, A., *English for Specific Purposes, A learning-centred approach*, Cambridge University Press, Seventh printing 1992, p. 147.
- (2) Heaton, J. B., *Classroom Testing*, Longman, Fourth impression, 1991, p. 30.
- (3) Purpura, J. E., *Assessing Grammar*, Cambridge University Press, Online Publication 2010, p.126.
- (4) Purpura, J. E., *Assessing Grammar*, Cambridge University Press, Online Publication 2010, p. 127.
- (5) Madsen, H. S., *Techniques in Testing*, Oxford University Press, First published 1983, p. 23.

BIBLIOGRAPHY

- [1] Arbogast, Janet, MEd; Kafer, Lindsey, MA; Chen, Julianna, MA., **Adult Learning** 25.4 (Nov 2014): 126-133.
- [2] Barnett, Candace W., Teaching Evaluation Practices in Colleges and Schools of Pharmacy, Matthews, Hewitt W., **American Journal of Pharmaceutical Education** 73.6 (2009): 103.
- [3] Göl, Ö and Nafalski, A., Collaborative learning in engineering education. *Global J. of Engng.Educ*, 11, 2, 173-180 (2007).
- [4] Heaton, J. B., Classroom Testing, Longman, Fourth impression, 1991.
- [5] Hutchinson, T., Waters, A., English for Specific Purposes, A learning-centred approach, Cambridge University Press, Seventh printing 1992.
- [6] Hughes, Bob, Ed. D. Evaluation as a Collaborative Activity to Learn Content Knowledge in a Graduate Course,
- [7] Madsen, H. S., Techniques in Testing, Oxford University Press, First published 1983.
- [8] Purpura, J. E., Assessing Grammar, Cambridge University Press, Online Publication 2010.
- [9] Tratnik, A., *Key Issues in Testing English for Specific Purposes*, ScriptaManent. 4(1) 3-13, Published by SDUTSJ. 2008.