STUDENTS AND INSTRUCTORS’ PERCEPTIONS OF OBJECTIVE TESTS USED TO ASSESS LANGUAGE PERFORMANCE AT UNIVERSITY LEVEL

(ÜNİVERSİTE SEVIYESİNDE DİL PERFORMANSINI DEĞERLENDİRMEDE KULLANILAN OBJEKTİF TEST TEKNİKLERİNE İLİŞKİN ÖĞRENCİ VE ÖĞRETMENLERİN ALGILAMALARI)

Dinçay KÖKSAL
Kürşat CESUR

ABSTRACT
The main purpose of this study is to have an idea on the students and instructors’ perceptions of the exams of Compulsory English Language Course which students take three hours a week at elementary level of English at Çanakkale Onsekiz Mart University. This study will also provide some suggestions for constructing and administering better language tests for the students taking Compulsory English Language Courses at different universities. According to the results of the study, the students and the instructors prefer other test techniques to be used in the exams in addition to the multiple-choice test technique. Results also reveal that instructors should also use teacher-made achievement tests in addition to the Standardized achievement tests as students have different language levels. In sum, the study presents the students and instructors’ perceptions of what has been done so far, what is being done now, and what can be done in the future to test students’ language performance at universities at elementary level.

Keywords: Testing, Objective Tests, Achievement Tests

ÖZ
Bu çalışmanın temel amacı, öğrencinin ve öğretmenin Çanakkale Onsekiz Mart Üniversitesinde haftada üç saat başlangıç seviyesinde aldıkları Zorunlu İngilizce Dersi sınavlarına ilişkin algılarını hakkında fikir sahibi olmaktır. Bu çalışma aynı zamanda başka üniversitelerde bu dersi alan öğrenciler için daha iyi testlerin geliştirilmesi ve uygulanması konusunda öneriler de sunacaktır. Çalışmanın sonuçlarına göre öğrenci ve öğretmenler çoktan seçmeli testlerle ek olarak başka test tekniklerini de kullanmayı tercih etmektedirler. Sonuçlar aynı zamanda öğrencilerin farklı seviyelerde olmasını dolayısıyla, öğretmenlerin merkezi sınav ek olarak kendi testleri tarafından hazırlanan testleri de kullanmalarını gerektiği ortaya koymaktadır. Özetle çalışma üniversitelerde öğrencilerin başlangıç seviyesinde dil performansının ölçülmesi için neler yapıldığı, neler yapıyormuş olduğu ve neler yapılabileceği hususunda öğrenci ve okutman algılamalarını göstermektedir.

Anahtar sözcükler: Test etme, Objektif Testler, Başarı Sınavları

1 This research is the part of MA thesis study.
2 Prof. Dr. Department of English Language Teaching, Faculty of Education, Çanakkale Onsekiz Mart University, Turkey. E-mail: dkoksal@yahoo.com
3 Instructor, School of Foreign Languages, Çanakkale Onsekiz Mart University, Turkey. E-mail: kcesurus@yahoo.com
INTRODUCTION

Testing is important for several reasons. To Madsen (1983:3), “testing is an important part of every teaching and learning experience”. It gives some ideas to the teachers for the future evaluation. Preparing accurate tests is important to get an accurate feedback on teaching. In the opinion of Hughes (1989:2-3) there are two main reasons for a test to be inaccurate. The first is about test content and techniques. For example, if the writing skill is only tested by multiple-choice items, the students practise such items rather than the skill of writing. Thus, the test becomes inaccurate. The second reason is the lack of reliability. To him, unreliability has two origins: “features of the test itself, and the way it is scored” (p. 3). This is the problem at Çanakkale Onsekiz Mart University (hereafter COMU) as the coursebook aims to include all language skills; however, students are tested through only multiple-choice test technique in their exams. The way the teachers score the tests at Compulsory English Language Courses can be totally same and reliable; however, features of the test and the technique (multiple-choice) used in these tests may not be accurate.

As the instructors and students’ perceptions of the test techniques were examined in this research, several studies helped the researchers to a great extent create their own questionnaires and determine the test techniques for objective testing of students’ language performance. Firstly, Dalyan’s (1990) study is a really good guide for this research as he suggested some techniques that can be scored objectively such as: multiple-choice, true-false, matching, rearrangement, addition, transformation, short-answer, and fill in the blank. According to the findings of his study, the most appropriate test items in the opinions of the teachers were respectively: multiple-choice, matching, true-false and fill in the blanks. He found out that most of the teachers always used multiple-choice test items to assess students’ language performances. Furthermore, in their study, Gelbal and Kelecioğlu (2007) found that teachers use traditional methods of testing as they know more about these methods than they do about alternative methods of testing. Similarly, Boud and Falchikov (2007) assert that multiple-choice items were in the past and still are generally recognized by students as the most widely applicable and useful means of testing.

In his research, Şahinel (1997) aimed to determine English lecturers’ opinions on the English language testing situations at the preparatory classes of Ankara University. He obtained the data by means of a questionnaire and he found that the test techniques used at the exams had not been arranged in the order of difficulty and the objectives of English language curriculum had not been taken into consideration by testers while constructing their tests.
addition to these studies; Aksan (2001) at Niğde University, Ösken (1999) at Hacettepe University and Serpil (2000) at Anadolu University administered questionnaires to the instructors and found out their perceptions of the content validity of the English language tests. The questions they asked to the instructors in their questionnaires shed light on constructing the questionnaires of this study.

Cohen (in Celce-Muricia, 2001:515) claims that students and teachers are afraid of the word “testing”. Students are afraid of tests because they think that they will not perform well. As for the teachers, “they do not construct tests and are not altogether satisfied with the results when they do. They are also suspicious of the standardized ... tests because they are not always sure what these tests are actually trying to measure” (Cohen, in Celce-Muricia, 2001:515). His claim and the literature reviewed raise important questions to be answered: is this the same case at COMU, what do the instructors think about the Standardized Achievement Tests (hereafter SATs) prepared by testers and are they satisfied with the practices of the testing office?

In the light of the literature discussed above, this study aims to find out the instructors and the students’ perceptions of the Compulsory English Language Course exams prepared by the testers of the testing office and to determine what other test techniques can also be used while evaluating students’ language performances at university level.

To achieve its aims, the study was guided by the following research questions (RQ):

**RQ1** Do the students and the instructors prefer the Compulsory English Language Course exams to be prepared by the testing office or by the instructor of the course?

- What is their most common reason for preferring either testing office or the instructor for the preparation and organization process of the exam?

**RQ2** What do the instructors and the students think about the exams that have been prepared by the test constructors of the testing office so far?

- Is there a significant difference among students’ thoughts about the exams prepared by the testing office in terms of their success?
- Is there a significant difference among instructors’ thoughts about the exams prepared by the testing office in terms of their job experience?

**RQ3** What other objective test techniques can be used in the mid-term and final exams of the course in the future according to students and instructors’ ideas?
METHODOLOGY

Quantitative research methodology was used in this descriptive study. Being a sub-category of a survey method, a questionnaire having some items and open-ended questions was prepared by the researchers.

Instruments

After reviewing the literature, the items of the questionnaire were constructed. Before applying the questionnaire for the piloting, the ideas of three experts on ‘English Language Teaching’ and one expert on ‘Measurement and Evaluation’ were asked. In the light of the experts’ ideas, some of the items were changed and some others were omitted. Also, some corrections were made in the questionnaire by the experts. Then, different parts of the questionnaire were organized in order to find answers to each research question of the study. The questionnaire conducted to the instructors and the students were the same. Only the items of the third part were organized for the students and the students’ questionnaire was in Turkish.

The first part of the questionnaire asked for the personal information about the participants. In the second part, whether the testing office or the instructor himself/herself should organize the exams of the Compulsory English Language Course was asked. The students and the instructors were required to specify the reasons why they prefer either testing office or the instructor to organize the exam in an open-ended question form. In the third part of the questionnaire, 5-point Likert-type Scale from ‘5=Strongly Agree’ to ‘1=Strongly Disagree’ was used in order to find out the students and instructors’ perceptions of the exams having been prepared by the testing office so far. Having had 20 items before the piloting, the third part of the questionnaire had 12 items in the end. Each of these items aims to get idea on one quality of the ‘Compulsory English Language Course’ exams. Finally, fourteen objective test techniques are presented in the fourth part of the questionnaire. In this part, students and instructors were asked which of these techniques could be used in centrally administered achievement tests. They were supposed to express their ideas on which techniques could be used in the exams by putting a cross (x) for the most appropriate test technique.

Having constructed all parts of the questionnaire, the researchers conducted the pilot study with 33 instructors and 121 students. Pilot study helped the researchers make necessary changes in the questionnaire for the main study. Researchers corrected the misspelled words and omitted the unnecessary ones for the main study. All in all, students, instructors and some experts’ ideas and comments let the researchers revise some of the items and make necessary changes for the main study.
Setting and Participants

The main study was conducted at COMU in February-March, 2008. The questionnaire was conducted at different departments of eight faculties and a college. Target population of the study consists of 34 instructors and 2928 second year students as all the students took the course in their first years and have ideas about the testing system. The number of the students was taken from the Students’ Affairs Office and that of the instructors was taken from the School of Foreign Languages. In a Compulsory English Language Course, students are aimed to be taught English at A1 level. The syllabus of the course is organized accordingly and the exams of the course are prepared by a testing office whose members are the instructors who are giving this course. They organize a centrally administered achievement test which consists of multiple-choice questions and is administered to all the first year students taking this course.

Instructors’ questionnaire was administered to 33 instructors. Most of the instructors attended a BA and MA Degree Courses on ‘Testing and Evaluation’ in ELT departments. Also, some instructors stated that they had participated in Seminars, Conferences, Symposiums, In-service Training Programmes, Workshops, and Courses by British Council on Testing and Evaluation. That is to say, they are knowledgeable enough to answer the questionnaire of this study efficiently.

As for the students, they were all in their second years at the university. They all had the Compulsory English Language Course the year before the study was conducted. Therefore, they were assumed to answer the questions in the questionnaire without any fear of their instructors or the possibility of the changes in the testing system. The researchers got the total population of the students (2928) from the Students’ Affairs Office. As it was almost impossible for the researchers to reach the whole population of the students (2928), the researchers made a Stratified Random Sampling. As suggested by Anderson (1990; cited in Balcı, 2005:95), with the ‘Confidence Level’ of 95% and with the ‘Confidence Interval’ of 5%; in a population consisting of 5000 members, 356 of these members can be used as a sample while carrying out a research. Thus, the researchers used 0.125 of the total population as a sample of this study which is 367.

Limitations of the Study

This research is limited to the opinions of 33 instructors and 367 students at COMU. The findings may not be generalized at an international level. However, they may reflect some important implications for other universities having SATs to assess their students’ language performances.

This study is also limited with the questions asked in the questionnaire. It may not reflect all other opinions of each student or that of each instructor.
Using only a questionnaire can be considered as another limitation. However, the number of the students to be questioned, the difficulty of reaching the instructors as they work in different faculties or colleges and the limited time to carry out this research forced the researcher to use a practical method for data collection, which is ‘questionnaire’.

**Data Collection and Analysis Procedures**

The researchers sent the instructors the questionnaire via e-mails. Data collection process from the instructors nearly took 20 days. As for the data collection process from the students, the researchers aimed to conduct the study to 367 students. In the first and second weeks of the spring term, the researchers conducted the questionnaires with the help of the lectures working in different faculties or college.

The data obtained through the questionnaire were analyzed via “Descriptive Statistics”, “One-way ANOVA”, and “Nonparametric Kruskal-Wallis Test” by using SPSS (Statistical Package for Social Sciences) 13.0 for Windows. Moreover, the researchers analyzed ‘the internal consistency’ of the items of the third part of questionnaire. Twelve items given in the third parts of the questionnaires are shown to have a high degree of internal consistency with values .84 for students’ and .86 for the instructors’ questionnaires. These values are generally acceptable for the data to be analyzed (Büyüköztürk, 2006).

**FINDINGS AND DISCUSSION**

In this part, the findings of the statistical analysis of the data are presented below the research questions. Then, discussions are made below each research question to be answered.

**RQ1** Do the students and the instructors prefer the Compulsory English Language Course exams to be prepared by the testing office or by the instructor of the course?

It was found out that nearly 73% of the instructors (n=24) want the exam to be prepared by a testing office. Only nine of the instructors (nearly 27% of the total population) prefer preparing the exam by themselves. Most of the instructors seem to be satisfied with the testing office according to these results or have some other reasons to prefer testing office. Unlike the instructors, more than half of (65.7%) the students (n=241) want the exam to be prepared by the instructor of the course. 34.3 percent of the students (n=126) prefer testing office for the organization process of the Compulsory English Language Course Exams. While instructors prefer testing office, students prefer the instructor of the course for the organization of the exams.
Their most common reasons for preferring either testing office or the instructor of the course will be discussed below.

**RQ1-A** What is their most common reason for preferring either testing office or the instructor for the preparation and organization process of the exam?

Among the 24 instructors who believe that testing office should organize the exams, the most common reason to prefer testing office is that ‘all instructors should follow the same curriculum’. They believe that if the exams were not prepared by the testing office, the instructors would not follow a common curriculum. To them, following a common curriculum is necessary to control and standardize their works on both teaching and testing.

Moreover, among the nine instructors who believe that the instructor of the course should organize the exams, the most common reason for the instructors’ preference is that “they want to use extra language teaching materials and prepare their own exam accordingly”. In the light of these findings, it seems that following a common curriculum and using extra materials in their teaching practices are considered to be important for the instructors.

126 students believe that testing office should organize the exams. Their most common reason to prefer testing office is that ‘centrally administered achievement tests hinder the instructors’ possible negative attitudes towards their learners’. Students believe that if the exams were not prepared by the testing office, the instructors would reflect their possible negative attitudes to the students’ grades or to the exam questions.

As discussed before, more than half of (65.7%) the students want the exam to be prepared by the instructor of the course. Their most common reason for preferring the instructor of the course is that ‘the instructors prepare their questions by taking their students’ different language levels into consideration’. Students want the exam to be prepared according to their levels. Although most of the students want the exams to be prepared by the instructor of the course, they have some doubts about those exams’ reliability.

**RQ2** What do the instructors and the students think about the exams that have been prepared by the test constructors of testing office so far? (See the Third Part of the Questionnaire for the items)

Total mean of the instructors’ perceptions (\(\bar{X}_{\text{Total}} = 3.46\)) shows that they are not much sure about the efficiency of the exams’ different characteristics. Although most of the instructors prefer the exam to be prepared by the testing office, they do not seem to be very satisfied with its current practices. Instructors agree that ‘the questions are clear enough to understand with the highest mean value of 4.18. Furthermore, they believe that the content of the questions matches the content they teach (\(\bar{X}_{\text{Item 2}} = 4.00\)) and the questions
represent the topic that they teach in the classroom ($\bar{x}_{item 4} = 4.00$). This shows that instructors agree that the exams have content validity. However, they do not agree that language is tested in the way it is taught in the exams ($\bar{x}_{item 8} = 2.91$). This may be because the instructors use many question types while teaching English but only multiple-choice while testing what they have taught. Hence, using other test techniques while assessing our learners’ language performance will help the testers construct effective tests. Mean value of the last item ($\bar{x}_{item 12} = 2.52$) confirms this result. That is to say, instructors prefer using other test techniques in addition to the multiple-choice test technique.

As for the students, they are undecided about the statements regarding the exams prepared by the testing office so far ($\bar{x}_{total} = 3.38$). Like the instructors, students also agree that the content of the questions matches the content they learned in the classroom ($\bar{x}_{item 2} = 3.72$) and the questions which are used in the exams match the course objectives ($\bar{x}_{item 1} = 3.71$). However, they do not agree that multiple-choice questions are efficient ($\bar{x}_{item 6} = 2.87$) and successful ($\bar{x}_{item 5} = 3.00$) in assessing their language performance as these items have the lowest mean values respectively. Therefore, using other test techniques in SATs of COMU is possible according to students and instructors’ ideas.

**RQ2-A** Is there a significant difference among students’ thoughts about the exams prepared by the testing office in terms of their success?

Among 367 students, 57 of them failed the compulsory English language course (0-59 Points – unsuccessful). 121 of them were successful enough to pass the course (60-79 points – successful). Finally, 188 of them were very successful in the course (80-100 Points – very successful). Mean values of the items where significant differences were found among the groups display that the more successful the students are, the more satisfied they are with the exams.

**RQ2-B** Is there a significant difference among instructors’ thoughts about the exams prepared by the testing office in terms of their job experience?

Among the twelve items evaluated, only the *sixth item* proved to have a significant difference between instructors’ perceptions of the exams and their job experience [$X^2_{(2)} = 7.276, p<.05$]. The findings suggest that as the instructors’ job experience increase, the mean rank of the item also increases. That is to say, the more experienced they are, the more efficient they find the exams prepared by the testing office. This can be because they do not want to prepare their own exams and they want to use readily made SATs.

**RQ3** What other objective test techniques can be used in the mid-term and final exams of the course in the future according to students and instructors’ ideas?
According to the instructors, multiple-choice ($\bar{x} = 4.36$), matching ($\bar{x} = 4.06$), ordering tasks ($\bar{x} = 4.03$), completion ($\bar{x} = 3.91$), short answer ($\bar{x} = 3.79$), error correction ($\bar{x} = 3.76$) and word changing ($\bar{x} = 3.67$) are among the most preferred objective test techniques. Instructors are undecided about using other techniques in the SATs of Compulsory English Language Courses.

As for the students; multiple-choice ($\bar{x} = 4.31$), short answer ($\bar{x} = 3.97$), true false ($\bar{x} = 3.92$), completion ($\bar{x} = 3.86$), matching ($\bar{x} = 3.78$), ordering tasks ($\bar{x} = 3.72$), word changing ($\bar{x} = 3.72$) and error correction ($\bar{x} = 3.59$) are among the most preferred objective test techniques. Students are also undecided about using other techniques in SATs as they are not much familiar with these techniques.

It is a surprisingly important finding that although their order of preferring objective test techniques slightly differs, the first eight test techniques that the students and instructors prefer to be used in the exams are completely same. Thus, these techniques should also be used at SATs.

**CONCLUSION AND SUGGESTIONS**

Most of the instructors prefer the exams to be prepared by the testers of the testing office. However, more than the half of the students prefer the instructor of the course to the testers of the testing office. Instructors’ most common reason to prefer testing office is that they believe they should all follow the same curriculum. As for the students, they prefer the instructors of the course because they commonly think that the instructors should prepare their own exams by taking their students’ different levels into consideration. Answers to the second research question show that both the instructors and the students have some doubts about the efficiency of the testing office’s current practices. Students’ ideas on the exams different characteristics differ in terms of their success. The more successful the students are, the more satisfied they are with the exams’ different characteristics. As for the instructors; it has been found that the more experienced they are, the more efficient they find the exams prepared by the testing office.

According to the results obtained from the answers to the third research question, instructors respectively prefer (1) multiple-choice questions, (2) matching, (3) ordering tasks, (4) completion, (5) true-false questions, (6) short-answer questions, (7) error correction and (8) word changing to be used in the SATs.

Bearing these conclusions in mind, some valuable suggestions can be made for effective testing of the students’ language performance. Firstly, in addition to the SATs of Compulsory English Language Course, instructors should also use teacher-made achievement tests as students have different
language levels. As Burke (2005:33) claims, “neither standardized tests alone, nor teacher[-made] assessments alone can provide a true picture of students’ learning”. Thus, instructors should also take the scores that students get from teacher-made achievement tests into consideration while assessing their learners’ language performance.

According to the results of the study, the students and the instructors prefer other test techniques to be used in SATs. Therefore, testers can include the eight objective test techniques, which were commonly preferred both by the instructors and the students, into the SATs. As H. D. Brown (2001) believes, a test is a method which consists of different techniques, procedures and items. Similarly, in the opinion of Gordon (1998:11), “it is important to understand … that no single assessment method can completely measure a student's range of skills and knowledge… Thus, it is necessary to use several types of assessment methods to help students learn…” Thus, Centrally Administered Achievement test of Compulsory English Language Course cannot be carried out using only multiple-choice test items. Not only these objective techniques, but also portfolio assessment can be a good source for assessing students’ language performances as it helps students discover and understand their strengths and weaknesses.

REFERENCES


Dalyan, A. (1990). The comments of EFL teachers on preparing achievement tests and an analysis of a sample test: A study at the engineering
APPENDIX

Dear Colleague,

The main purpose of this questionnaire is to find out the instructors’ perceptions of the Compulsory English Language Course exams. Your answers will provide important information about the effectiveness of the testing practices at Compulsory English Language Courses. The answers to this questionnaire will be kept confidential. Thank you for your kind cooperation in completing this questionnaire.

Part 1 for the Students

BÖLÜM -1: Sizin için uygun olan seçeneğe çarpı (X) işaretini koyunuz.

Cinsiyet:  a. Bay ( )  b. Bayan ( )

Birinci sınavdaki “Zorunlu İngilizce Dersi” Bahar Dönemi harf notunuz:

a. AA ( )  b. BA ( )  c. BB ( )  d. CB ( )  e. CC ( )

f. DC ( )  g. DD ( )  h. FD ( )  i. FF ( )  j. DS ( )

Eğitim gördüğünüz Fakülte ya da Yüksek Okul:


d. Mimarlık Mühendislik F. ( )  e. İlahiyat F. ( )  f. Biga İktisadi ve İdari Bil. F. ( )

g. Su Ürünleri F. ( )  h. Ziraat F. ( )  i. Sağlık YO ( )

Part 1 for the Instructors

PART -1: Please put a cross (X) into the brackets which is appropriate for you. If there is another choice, please specify it into the “other” section.

Teaching Experience:  a. 0-3 years ( )  b. 4-6 years ( )  c. 7 years and more

Graduation:  a. BA Degree ( )  b. MA Degree ( )

Department:

a. English Language Teaching:  BA ( ) / MA ( )  b. English Language and Literature:  BA ( ) / MA ( )

c. English Linguistics:  BA ( ) / MA ( )  d. Translation and Interpretation:  BA ( ) / MA ( )

e. American Culture and Literature:  BA ( ) / MA ( )  f. Other ( )

Have you attended to the things below on “Testing and Evaluation”? (More than one option is possible)

a. BA degree course ( )  b. MA degree course ( )  c. Seminar ( )

d. Conference ( )  e. Symposium ( )  f. Others ( )

PART -2: Please put a cross (X) into the box which is appropriate for you.

Do you think the exams of the “Compulsory English Language Course” should be organized by the testing office or by the instructor of the course himself/herself? Who should organize the exams of the course?

1. Testing office: ( )  2. Instructor himself/herself: ( )

Why do you prefer either testing office or the instructor for the preparation and organization process of the exam?

………………………………………………………………………………………………………………

………………………………………………………………………………………………………………

………………………………………………………………………………………………………………

………………………………………………………………………………………………………………

………………………………………………………………………………………………………………

………………………………………………………………………………………………………………

………………………………………………………………………………………………………………

………………………………………………………………………………………………………………

………………………………………………………………………………………………………………
### PART -4-: Please put a cross (X) into the box which is appropriate for you. Which of these techniques can be used in centrally administered achievement tests?

<table>
<thead>
<tr>
<th>OBJECTIVE TEST TECHNIQUES WHICH</th>
<th>CAN BE USED IN OUR EXAMS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>STRONGLY AGREE</td>
</tr>
<tr>
<td>1. Multiple-Choice Questions</td>
<td></td>
</tr>
<tr>
<td>2. Short-Answer Questions</td>
<td></td>
</tr>
<tr>
<td>3. True-False Questions</td>
<td></td>
</tr>
<tr>
<td>4. Matching</td>
<td></td>
</tr>
<tr>
<td>5. Completion</td>
<td></td>
</tr>
<tr>
<td>6. Cloze Test</td>
<td></td>
</tr>
<tr>
<td>7. C-Test</td>
<td></td>
</tr>
<tr>
<td>8. Cloze Elide Test</td>
<td></td>
</tr>
<tr>
<td>9. Ordering Tasks (Rearrangement)</td>
<td></td>
</tr>
<tr>
<td>10. Error Correction</td>
<td></td>
</tr>
<tr>
<td>11. Transformation</td>
<td></td>
</tr>
<tr>
<td>12. Combination</td>
<td></td>
</tr>
<tr>
<td>13. Addition</td>
<td></td>
</tr>
<tr>
<td>14. Word Changing</td>
<td></td>
</tr>
</tbody>
</table>

Please specify below if there are any other techniques you would like to add.

15.

16.