

**PERCEPTIONS OF STUDENTS TOWARDS THE IMPLEMENTATION OF
PERFORMANCE-BASED ASSESSMENT IN VOCATIONAL ENGLISH
COURSE**



Dilara BİDAV

JANUARY 2021

**PERCEPTIONS OF STUDENTS TOWARDS THE IMPLEMENTATION OF
PERFORMANCE-BASED ASSESSMENT IN VOCATIONAL ENGLISH
COURSE**

**A THESIS SUBMITTED TO THE
GRADUATE SCHOOL PROGRAMS
OF
BAHCESEHIR UNIVERSITY**

BY

DİLARA BİDAV

**IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF ARTS
IN THE DEPARTMENT OF ENGLISH LANGUAGE TEACHING**

JANUARY 2021



**T.C.
BAHCESEHIR UNIVERSITY
GRADUATE SCHOOL**

22/ 01/ 2021

MASTER THESIS APPROVAL FORM

Program Name:	English Language Education
Student's Name and Surname:	Dilara Bidav
Name Of The Thesis:	Perceptions of students towards the implementation of performance-based assessment in vocational English course
Thesis Defense Date:	22. 01. 2021

This thesis has been approved by the Graduate School which has fulfilled the necessary conditions as Master thesis.

Assoc. Prof. Dr. Burak KÜNTAY
Institute Director

This thesis was read by us, quality and content as a Master's thesis has been seen and accepted as sufficient.

	Title/Name	Signature
Thesis Advisor's	Prof. Dr. Kenan Dikilitaş	
Member's	Dr. Aslı Sağlam	
Member's	Dr. Özlem Yalçın Çolakoğlu	



I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last Name : Dilara BİDAV

Signature :

ABSTRACT

PERCEPTIONS OF STUDENTS TOWARDS THE IMPLEMENTATION OF PERFORMANCE- BASED ASSESSMENT IN VOCATIONAL ENGLISH COURSE

Bidav, Dilara

Master's Thesis, English Language Education

Supervisor: Prof. Dr. Kenan DİKİLİTAŞ

January 2021, 75 pages

This thesis aims to investigate the student perceptions towards the implementation of performance- based assessment (PBA) in vocational English course. In the light of this purpose, the study was conducted on 52 students studying in a vocational and technical high school in Izmir. Mix research design attributes used in order to obtain data. Based on this information, the quantitative data was collected by using a 5 point Likert scale in this study. Additively, the study made use of qualitative data which were obtained from the students' responses to the focus group discussion for the purpose of gaining insights into students' perceptions on PBA. This paper provides guidelines for English for Specific Purposes (ESP) learners' attitudes on the alternative assessment with a specific focus on vocational English courses. Findings for the current study suggest that most of the students are in favor of performance- based assessment tasks (PT). The collaborative nature of the focus group mostly offers some benefits of performance- based assessmet tasks; however, it has some challenges.

Keywords: Performance- based Assessment, ESP, Benefits, Challenges, Private Vocational and Technical High School

ÖZ

PERFORMANSA DAYALI ÖLÇME DEĞERLENDİRMENİN TEKNİK İNGİLİZCE DERSİNDE UYGULANMASINA DAİR ÖĞRENCİ TUTUMU

Bidav, Dilara

Yüksek Lisans, İngiliz Dili ve Eğitimi Yüksek Lisans Programı

Tez Yöneticisi: Prof. Dr. Kenan DİKİLİTAŞ

Ocak 2021, 75 sayfa

Bu çalışmanın amacı performansa dayalı ölçme değerlendirme teknik İngilizce dersinde uygulanmasına dair öğrenci tutumunu irdelemektir. Bu amaç doğrultusunda Mesleki ve Teknik Anadolu Lisesinde okuyan 52 öğrenci ile çalışma yürütülmüştür. Nitel ve nicel veriler anket ve öğrencilerle yapılan odak grup görüşmeleri ile elde edilmiştir. Çalışmanın bulguları öğrencilerin performansa dayalı ölçme değerlendirmeye karşı pozitif düşüncelerini ifade ederken aynı zamanda negative sonuçları olduğunu da belirtmiştir. Bu çalışma öğrencilerin mesleki İngilizce derslerinde alternatif değerlendirmeye yönelik tutumlarına dair önemli bulgular ortaya koymuştur. Öğrenciler performansa dayalı ölçme değerlendirme teknik İngilizce dersinde kullanılması konusunda ortak bir kanıda olup nitel ve nicel veri analizleri de bunu desteklemektedir.

Anahtar Kelimeler: Performansa Dayalı Ölçme Değerlendirme, Özel Amaçlı İngilizce, Faydalar, Zorluklar, Mesleki ve Teknik Anadolu Lisesi



To the ones who stood by my side through it all...

ACKNOWLEDGMENTS

Throughout this study, I have been luckily surrounded by endless support and guidance. I would like to express my deepest feelings of gratitude to those people. This accomplishment wouldn't have been possible without any of them.

First, I deeply acknowledge my supervisor, Professor Dr. Kenan Dikilitaş, for his guidance, assistance and constructive suggestions.

I would also like to express my gratitude to my committee members, Dr. Aslı Sağlam, and Dr. Özlem Yalçın Çolakoğlu for their insightful comments and suggestions.

I am grateful to all of my friends who were the real witnesses of my pain and joy. Dicle Ay Karaoğlu, with whom I started this journey, encouraged me anytime I felt demotivated. I couldn't have dared to do this without her support and motivation. I must express my very profound gratitude to my dearest friends, Çiğdem Uluçay, Çiçek Hızlı, Tuğçe Karaulutaş and Merve Demir, for their unfailing support and continuous encouragement.

I owe my deepest gratitude and thanks to Gamze Maden and Sercan Çelik, they are such true friends anyone could wish for. Without their help, this work would not be possible as they were all by my side when I questioned my patience, strength and hope.

Last but not the least, I would like to express my sincere love and gratitude to my family, Gülay Bidav, Şahin Bidav, and Mustafa Bidav for providing me with unfailing support and continuous encouragement throughout my years of study.

TABLE OF CONTENTS

ETHICAL CONDUCT.....	iii
ABSTRACT.....	iv
ÖZ.....	v
DEDICATION.....	vi
ACKNOWLEDGMENTS.....	vii
TABLE OF CONTENTS.....	viii
LIST OF TABLES.....	x
LIST OF FIGURES/ILLUSTRATIONS/SCHEMES.....	xi
LIST OF SYMBOLS/ABBREVIATIONS.....	xii
Chapter 1: Introduction.....	1
1.1 Background of the Study.....	1
1.2 Statement of the Problem.....	3
1.3 Purpose of the Study.....	4
1.4 Research Questions.....	4
1.5 Significance of the Study.....	5
1.6 Definitions.....	5
Chapter 2: Literature Review.....	7
2.1 ESP.....	7
2.1.1 Aviation English.....	10
2.2 Language Assessment.....	11
2.2.1 Traditional Assessment.....	12
2.2.2 Alternative Assessment.....	14
2.2.3 Traditional Assessment vs Alternative Assessment.....	16
2.2.4 Performance- Based Assessment.....	18
Chapter 3: Methodology.....	21
3.1 Research Design.....	21

3.2	Setting and Participants	22
3.2.1	Setting	22
3.2.2	Participants	24
3.3	Procedures	25
3.3.1	Data Collection Instruments.....	25
3.3.1.1	Questionnaire	25
3.3.1.2	Focus Group Discussion	26
3.3.2	Data Collection Procedures	27
3.3.2.1	PT Implementation Process	27
3.3.2.2	Grading and Scoring Rubrics	31
3.3.3	Data Analysis Procedures	31
3.3.4	Reliability and Validity.....	32
3.3.5	Trustworthiness.....	33
3.4	Limitations and Delimitations.....	34
Chapter 4:	Findings	36
4.1	Findings of Research Question 1	36
4.2	Findings of Research Question 2	42
4.3	Findings of Research Question 3	48
Chapter 5:	Discussion and Conclusions	50
5.1	Discussion of Findings for Research Question 1	50
5.2	Discussion of Findings for Research Question 2	51
5.3	Discussion of Findings for Research Question 3	53
5.4	Conclusions	54
5.5	Recommendations	55
REFERENCES	56
APPENDICES	63
A.	Consent Form	63
B.	Questionnaire	64
C.	Scoring Rubric of Writing PT	68
D.	Scoring Rubric of Vocabulary PT	70
E.	Scoring Rubric of Speaking PT	71
F.	Task Specifications	72

G. Curriculum Vitae.....	75
--------------------------	----

LIST OF TABLES

TABLES

Table 1 Differences Between Traditional and Alternative Assessment.....	18
Table 2 Demographic Data of the Participants of the Study.....	24
Table 3 Weekly Schedule of the PT Implementation	28
Table 4 Cronbach’s Alpha	33
Table 5 Reliability Criteria for <i>a</i> Value	33
Table 6 Descriptive Statistics Results of Students’ Ideas Related to Phases of the Performance- based Assessment.....	36
Table 7 Perceptions of the Students on the Items Related to Task Familiarity Phase	37
Table 8 Perceptions of the Students on the Items Related to Application Phase...	38
Table 9 Perceptions of the Students on the Items Related to Scoring Phase	39
Table 10 Perceptions of the Students on the Items Related to Learning Outcomes Phase	40
Table 11 Perceptions of the Students on the Items Related to Program Consistency	41
Table 12 Perceptions of the Students towards PBA according to Their Proficiency Level	48

LIST OF FIGURES

FIGURES

Figure 1 The Tree of ELT	9
Figure 2 Grading system of vocational English class	31

LIST OF ABBREVIATIONS

ELT	English Language Teaching
ESP	English for Specific Purposes
GE	General English
MoNE	Ministry of National Education
FGD	Focus Group Discussion
PBA	Performance-based Assessment
PT	Performance- based Assessment Task
TA	Traditional Assessment
AMT	Aircraft Maintenance Technician
SPSS	Statistical Package for Social Sciences
KMO	Kaiser-Mayer-Olkin
ATC	Air Traffic Controller

Chapter 1

Introduction

The aim of this chapter is to provide readers with an outline of the present study. It starts with a brief description of background of the study which continues with the statement of the problem, purpose of the study, research questions, significance of the study and key definitions respectively.

1.1 Background of the Study

Assessment is integrated in every part of the language teaching and learning process. Brown (2004) states that whenever a student responds to a question, offers a comment, or tries out a new word or structure, the teacher subconsciously makes an assessment of the student's performance. Nowadays, ESP testing and assessment have gained prominence in many institutions which offer ESP courses. Basturkmen and Elder (2006) draw our attention to two practical arguments for ESP testing. According to them, ESP tests make the methods, goals, and outcomes of assessment more transparent, and hence more convincing to end users and ESP tests are more likely to generate. For this reason, users will be more convinced and also ESP tests will generate teaching activity which is seen by learners as relevant to their needs. Benavent and Peñamaría (2011) support this argument by saying that this means that the backwash (the effect on the learning process) achieved by a finely-tuned ESP test design will be beneficial. According to Stoica (2006), the focus of ESP is that "English is not taught as a subject separated from the students' real world; instead it is integrated into a subject matter area important to the others"(p. 326). In other words, ESP assessment should be based on performing real-life tasks. Regarding this fact, it is inevitable to make use of new trends in assessment methods implemented in ESP classes.

Since 1990s, the researchers have been investigating and suggesting the innovative types of assessment methods in language teaching. McNamara (2001)

defines alternative assessment as a movement “away from the use of standardized multiple-choice tests in favor of more complex performance based assessments” (p.329) and Hancock (1994) described the term as “an ongoing process involving the student and teacher in making judgments about the student’s progress in language using non-conventional strategies” (p.3). In alternative assessment, learners acquire problem solving and higher level thinking skills, real-world contexts or simulations are utilised, and both process and products are focused on (Çakır, 2013). Additionally, Cirit (2014) highlights that alternative methods reflect students’ performance in educational context which helps the instructors collect information about their students' interests, talents, abilities, and potentials. Moreover, Hamayan (1995) adds that “Alternative assessment provides an insight into individual students’ language proficiency. The information obtained from alternative assessment is extensive and reflects a wide range of abilities and skills in language in a variety of contexts” (p.216).

On the contrary to alternative methods in assessment, many institutions mostly prefer traditional assessment (TA) tools which are product oriented. On the contrary, the balance between assessment of not only product but also process is crucial for the effective assessment. Çakır (2013) argues that in TA tests students are presented with language and required to pick the correct answer from among a limited set of options, no language is created by the students. Many scholars point out that traditional assessment approach is insufficient to reflect the students’ success or failure in detail. Balliro (1993) mentions the dissatisfaction with the traditional assessment methods by stating that the traditional assessment methods remain incapable of sufficiently representing the learner strengths and true progress. To this end, Gökhan (2004) points out that in alternative assessment students use the target language in similar contexts and for the same purposes as they need to use it in real life. Brown and Hudson (1998) also indicate that the traditional assessment methods are not capable of reflecting the authentic language. Additively, Hamayan (1995) states that alternative assessment procedures are based on activities that have authentic communicative function.

As one of the alternative assessment methods, PBA “represents a set of strategies for the application of knowledge, skills, and work habits through the performance of tasks that are meaningful and engaging to students” (Yıldırım& Örsdemir, 2013, p.563). According to Stiggins (2001), PBA represents a set of strategies that asks students to use their knowledge and skills to create a product or perform a task that is authentic and meaningful to them based on certain predetermined criteria. For

instance, the tasks may include giving oral reports, writing essays, participating in interviews, among others. It is clear that PBA and evaluation methods should also be used in education processes in addition to traditional assessment tools due to many reasons such as inability to measure high- level skills and inability to reveal the mental schema of the student (Enger& Yager, 1998; Shepard, 2008).

Taking all points above into the consideration, this study aims to investigate the use of PTs in vocational English classes and its effectiveness from students' perspectives. Also, it aims to find out students' perception on the overall process consisting task familiarity, application, scoring, learning outcomes and consistency with the current vocational English program.

1.2 Statement of the Problem

Turkish Ministry of National Education (MoNE) offers vocational English courses in vocational high schools. TA methods are predominantly used in order to assess students' achievement in vocational English classes. According to Balliro (1993), the TA methods are not satisfactory for assessing students' success as they are not adequate for reflecting the learner's strengths and actual progress sufficiently.

In addition to that, Law and Eckes (1995) highlight a significant problem with TA by saying that they just assess students' lower order thinking skills. Recently, TA tools are challenged by alternative assessment methods. Gökhan (2004) states that alternative assessment should be authentic in ELT contexts as it engages students with meaningful and communicative tasks which reflect real-life situations. In the light of this information, as a requirement of ESP assessment Ardiyansah (2019) puts forward that the assessment tools implemented in ESP classes should be authentic and based on the real world activities.

With the aim of defining PBA as one of the alternative assessment methods, Shohamy (1995) states that “performance testing opens a new era of assessment where language is being assessed in context along with other skills, not in an isolated, decontextualized manner” (p.206). Hence, it was clear that using only TA methods to assess all skills and language elements in the target language was not sufficient enough as students needed to produce the language in the target context. In parallel with this problem, efficiency of assessment in vocational English courses comes under question.

With the aim of finding an acceptable solution to that, uses of PBA methods in addition to TA tools were investigated in this study as Quansah (2018) remarks:

“in vocational and technical schools, for example, performance based assessments are emphasised more than traditional assessment. It must be re-echoed that, it is essential for traditional assessment procedures to be used to assess whether learners can remember facts and knowledge learnt before alternative assessment used”(p. 23).

It is clear that there is a gap regarding the implementation of PBA as an assessment tool in vocational English courses offered in vocational and technical high schools. Hence, this study aims to investigate students' perceptions on the implementation of PBA in the context of aviation English. It also aims to find out students' attitude towards the whole implementation process of PTs' by focusing on phases such as task familiarity, application, scoring, learning outcomes and consistency with the vocational English program. Therefore, outcomes of the study may provide some insights for the teachers and the researchers studying in this field in order to fulfil this need in the literature.

1.3 Purpose of the Study

This study was implemented in a vocational and technical high school where traditional assessment practices were used predominantly in vocational English courses. The researcher combined traditional assessment tools and alternative assessment methods for learners' overall success. Therefore, this study aims to understand perceptions of the students concerning implementation of PTs as a supplementary to the TA in vocational English courses. In view of all that has been mentioned so far, it may also help to see strenghts and weaknesses of the implementation of PTs in vocational English courses.

1.4 Research Questions

Purpose of this study is to answer these questions:

1. What are the students' perceptions of the PTs regarding task familiarity, application, scoring, learning outcomes, and program consistency?
2. What are the impacts of PBA tasks on students?

3. Do the students' perceptions of PBA tasks vary based on their proficiency level?

1.5 Significance of the Study

This study aims to give a detailed insight on the vocational school students' attitudes towards the use of PBA tools in vocational English courses. The literature will be enriched by the outcomes of this study by shedding light on implementing efficient assessment tools through PTs. The findings hopefully intend to be helpful to teachers or researchers in finding ways to offer students with successful assessment strategies and in guiding them how to integrate the tasks into their teaching and assessment preferences.

1.6 Definitions

The key definitions of the significant terms of the study are given below.

Assessment: The act of judging or deciding the amount, value, quality, or importance of something, or the judgment or decision that is made (Online Cambridge Dictionary, 2020).

Alternative Assessment: It means measuring performance through a variety of creative, open-ended, or communicative means (Brown, 2004).

Traditional Assessment: It is a way of testing learner's achievement by asking some objective questions which contain one correct or best answer (Brown, 2004).

Performance- based Assessment: A performance- based assessment measures students' ability to apply the skills and knowledge learned from a unit or units of study. Typically, the task challenges students to use their higher- order thinking skills to create a product or complete a process (Chun, 2010).

English for Specific Purposes: It means teaching English to be used in a specific field of activity such as science and business (Online Cambridge Dictionary, 2020).

Chapter 2

Literature Review

The objective of this study is to search for students' attitude towards the use of PBA methods in vocational English course, namely aviation English course, at a vocational and technical high school in İzmir. This chapter gives a review of the literature on ESP, aviation English, traditional assessment, language assessment, alternative assessment, alternative vs. traditional assessment, and performance-based assessment. To put it briefly, this study aims to investigate some basic kinds of language testing, ways of assessing vocational English and previous studies focusing on PBA and student perceptions will be presented with the aim of providing in-depth information about assessment and its related concepts in a vocational and technical high school.

2.1 English for Specific Purposes (ESP)

As with many improvements in human activity, English for Specific Purposes (ESP) is a “phenomenon growing out of many converging trends instead of being planned and coherent movement” (Hutchinson & Waters, 1992). Demirdöken (2019) explains the reason of these trends by stating that ESP first appeared due to some factors caused by remarkable changes in the world including changes in the necessity of learning a foreign language, the need for the new world order, and learner-based education. ESP is defined as a sub-branch of English as a foreign language (EFL) or English as a second language (ESL) with the aim of teaching English to learners who want to be proficient in the target language in the target context (Demirdöken, 2019). ESP became a current issue when General English (GE) is not enough for students who have special goals to learn English. Rahman (2015) points out that ESP learners are familiar with English and learn the language for the purpose of performing specific profession-related activities and communicating a group of professional skills. In other words, the term ‘specific’ in ESP clearly identifies that all students who want to be proficient in ESP have a particular goal in order to learn English

because only GE cannot meet the needs of these students who converge to ESP. Learners of ESP are generally students who already have familiarity with English and learn it with the aim of being capable of performing job-related subjects such as science, tourism and medicine. On the other hand, Hutchinson and Waters (1987) define ESP as a way of teaching language where learners' reason for learning it is the basis of decisions about content and method. ESP can be defined as an approach that depends on learner demands. ESP provides learners with proficiency in the target language and allows them to learn necessary vocabulary terms related to the target domain. Stoica (2006) states that ESP focal point is teaching English which is not separated from students' real world. Also, it draws attention to importance of vocabulary items in ESP by saying that ESP is more like a combination of subject matter and English Language teaching. Therefore, putting great importance on the vocabulary and seeing it as central is considered as a target of an ESP course. In an agreement to Stoica, Demirdöken (2019) adds that the structure of ESP courses is not based on a grammar- based teaching; however, it focuses on the target language in the target context.

GE puts importance on four language skills equally such as speaking, reading, listening, and writing. However, Rahman (2015) asserts that ESP requires analysis in order to determine which language skills the learners mostly need, and also the course syllabus should be designed on the basis of the learners' need. Hence, designing that required syllabus depends on the students' job- related profession needs. For example, if a student wants to work in tourism, the ESP course which should be built on the basis of the student's needs and functions which concentrate on increasing the development of communication skills. Ahmed (2014) sums up ESP as a particular way of teaching language which needs to make decisions about content and method which are based on the learner's reason to learn the language.

The definition which tells us about absolute as well as variable aspects of ESP shows what ESP is and is not because it is a very broad term. Two different branches such as English for Occupational Purposes (EOP) and English for Academic Purposes (EAP) were established under the umbrella of ESP. EAP means any teaching related to requirements of academic study (Rahman, 2015)., whereas EOP

focuses on the language used in work or training. Hutchinson and Waters (1987) show the history of EAP and EOP by representing the relationship in the tree of ELT.

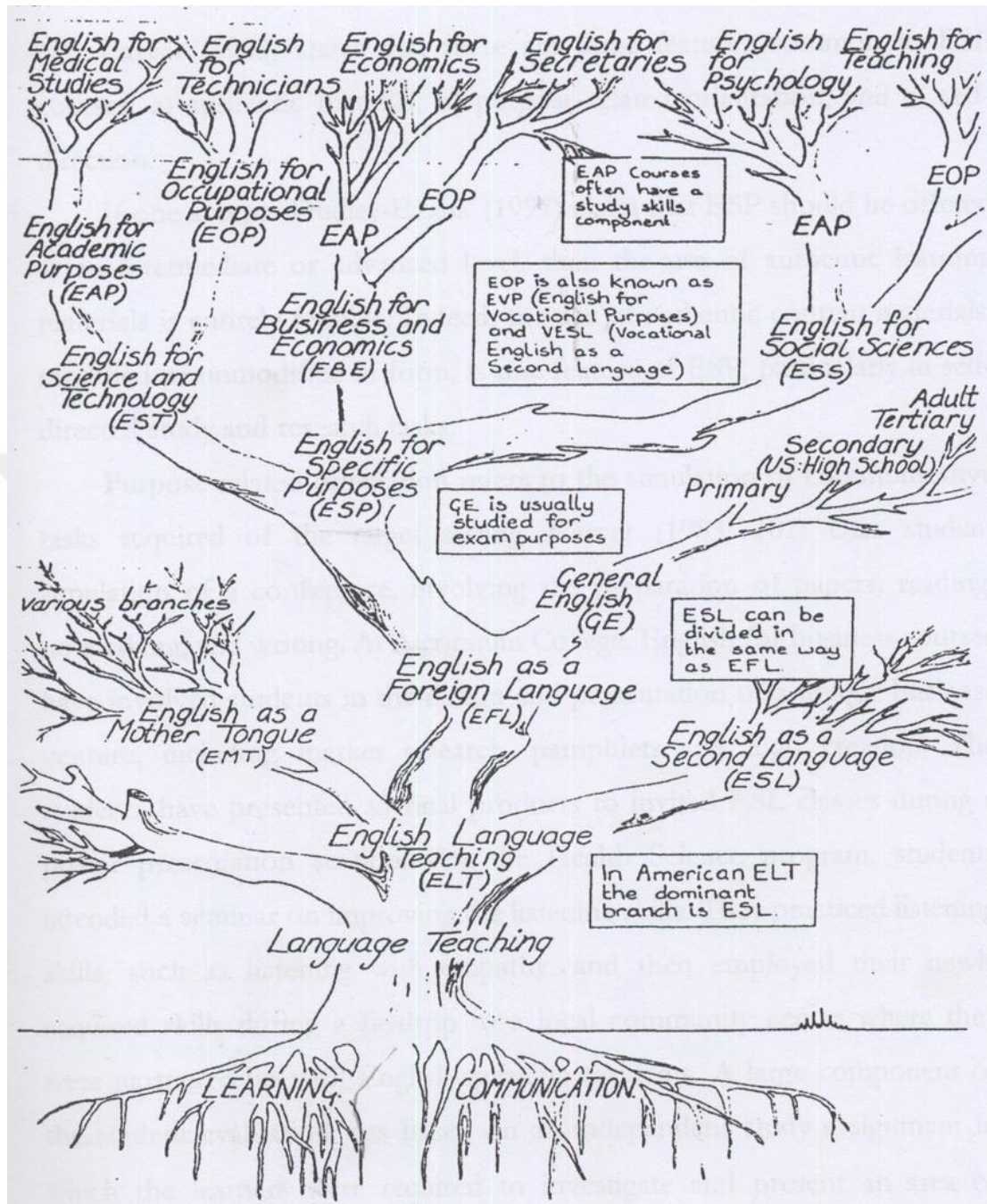


Figure 1. The tree of ELT. Adapted from *English for Specific Purposes* (p.17), by, T. Hutchinson & A. Waters, 1987, Cambridge: Cambridge University Press.

Hutchinson and Waters (1987) states that ELT tree shows that ESP is divided into three pieces;

- a) EST (English for Science and Technology)
- b) EBE (English for Business and Economics)
- c) ESS (English for Social Sciences).

Each subset of ESP is also broken down into two main branches: EAP and EOP. English for Psychology can be regarded an example of EAP while English for Technicians is an example of EOP. Hutchinson and Waters (1987) highlight a crucial point of this difference between EOP and EAP by stating that this is not a clear distinction: individuals can study and work at the same time. In many cases, the student will use the language he/she has learned in a study context when he/she starts working or returns to work. To put it briefly, ESP in the context of aviation English can be called as EOP.

2.1.1 Aviation English. With the aim of defining radio- telephony phraseology as well as English used by air traffic controllers, pilots, flight attendants, maintenance technicians, airport staff and managers in the aviation industry and so forth, a new and more general term Aviation English has been used since the 1970s. Just like other types of ESP, the aviation English used by the auxiliary and other aviation staff is adapted to the specific field and context, in addition to using the pronunciation, syntax, lexicon, etc. of the conventional English (Serpil, 2017). Although it seems that aviation English context refers to just the language used between Air Traffic Controllers and pilots, Aiguo (2008) claims that aviation English is not only limited to pilot and controller communications but also include using English related to any other aspect of aviation: pilots needs the language for announcements, briefings, and flight deck communication. Moreover, the maintenance technicians, managers, flight attendants, officials, and dispatchers use language in the aviation industry or even students use the English language in aviation or aeronautical universities. Serpil (2017) adds that although it contains the phraseology specified by the International Civil Aviation Organization (ICAO), aviation English should not be limited to this, in some cases, it may necessitate the use of GE. Some research focused on the role of terminology and language use for Aviation safety. Tan (2016) illustrates the importance of communication breakdown between maintenance technicians by giving an example of an accident caused because of lack of understanding of the terminologies:

“Avianca Flight 052 departed Medellin, Colombia, for New York City, USA, on January 25, 1990. It was delayed 77 minutes en route by bad weather and ran low on fuel. The crew notified the control tower that they were ‘just running out of fuel’ and needed priority, but they never used the technical terms that signaled an emergency: ‘minimum fuel’ or ‘emergency fuel’...Six minutes later, they radioed that they’d just lost two engines and needed priority. The controller cleared the plane to approach and asked if they had enough fuel to

land. There was no response. The plane crashed, killing 73 people on board, including the pilots” (p. 8).

As stated by the case study, the phrase running out of fuel would be broadly understood by many industries as well as the general public in everyday usage, but it does not necessarily trigger an emergency protocol in the aviation context (Tan, 2016).

When it comes to a study carried out by Eckert (1997), it focuses on using simplified English for the improvement of non-native technician students' task comprehension. Eckert (1997) highlights that most of the accidents that happened between 1979 and 1991 are because of the maintenance factors. Serpil (2017) explains this issue by stating that in the United States or in other foreign countries, if an airline company is not authorized to translate the manuals with the approval of Federal Aviation Academy (FAA), the FAA- certified technicians have to follow the manuals written in English. In Eckert's study, Public Broadcasting Service reported that being fluent in English was not a requirement for aircraft maintenance personnel in order to follow the instructions in (AMM) Aircraft Maintenance Manual. In an episode of a documentary released by PBS, a reporter had an interview with the aircraft maintenance technicians in Costa Rica and asked them to present the maintenance manual in the facility because during the interview it was very clear that they had lack of effective communication skills. Eckert (1997) highlights that it was evident that technicians were not fluent in English. Hence, improving speaking skills of aircraft maintenance personnel is crucial.

2.2 Assessment of Language

The phrases ‘evaluation’, ‘testing’ and ‘assessment’ are still commonly utilized in the literature similarly in order to allude to a variety of judgement and choices about a particular language program, and its components. They are quite different from each other. To begin with evaluation, it is defined as interpreting the assessment results and describing the benefits of a student's performance regarding learner's standards or expectations of performance (Cheng et al., 2004). Regarding the difference between assessment and testing, many people tend to think that they are synonymous words. However, they attribute various concepts. Testing is a procedure which is designed for administrative purposes and occurs periodically based on a curriculum. Additionally, assessment is a much broader and continual process (Brown, 2004). Law and Eckes

(1995) define testing as timed, single-occasion, and undimensional exercise typically in short-answer or multiple choice forms. On the other hand, in assessment, information is collected at different times and contexts using various ways (Law and Eckes, 1995). According to Bachman (1991), assessment means that the interesting information about something is collected based on the systematic and substantially grounded procedures. Additionally, Stoica (2006) illustrates that assessment is being aware of the students and the quality of their learning. Moreover, Dietel, Herman, and Knuth (1991) state that assessment is a type of a method so that teachers can figure out the present knowledge that students possess. Moreover, basic objectives of assessment are summed up by Stoica (2006) as follows: considering crucial skills and knowledge; checking the improvement in time; searching for the difficulties that students may face; analyzing teaching strategies and efficacy of the program; encouraging students to study. Researchers (Huerta- Macias, 1995; Brown & Hudson, 1998) argue that two important themes emerged as two main ways of assessment: traditional assessment and alternative assessment. As Brown (1998) states in the early 1990s, on the contrary to the notion that traditional tests should be used to measure all skills of people, a new concept appeared as ‘alternative’ assessment. Hamayan (1995) defines the main attributes of alternative assessment as: a holistic view of language, proximity to actual language use and performance, developmental appropriateness, an integrative view of learning, and multiple referencing. On the other hand, traditional testing is defined as a kind of test in which students work alone and tests are indirect and de-contextualized along with norm-referenced (Brown, 2004).

2.2.1 Traditional assessment. Every person has been tested with one of the types of traditional assessment tools once in their life. In traditional assessment, learner achievement is tested by asking objective questions containing the best or correct answer (Brown, 2004). In order to support Brown, Stocia (2006) highlights that students in traditional assessments should recognize a particular choice as the best answer to the question. As noted by Quansah (2018), traditional assessment means the conventional methods through which written documents, such as exams or quizzes are produced for testing. Traditional assessment tools are still the most commonly used in classrooms. Ozuslu (2018) supports this argument by saying that the most common assessment practices in classroom are the paper-pencil tests which are based on

different procedures such as administering quizzes, choosing the suitable assessment methods, revising any teacher-made test, or writing test items to match the instructional objectives.

A significant analysis and discussion on the traditional assessment was presented by Cumming (2009) by stating that although the pedagogical requirements are met by the innovative, criterion-based assessments in the educational systems, the norm-based language tests authority keep their role in educational field. As indicated by Brown and Hudson (1998), the traditional methods can be used by the teachers to assess students' understanding of a specific point. Thus, the outcomes of the traditional examinations show the teachers whether the objectives of a course were fulfilled or not (Brinke et al., 2007).

Weaknesses of the traditional methods of assessment were repeatedly voiced by various scholars. As argued by Brown and Hudson (1998), the authentic language could not be reflected by the traditional assessment methods. Nasab (2015) supports this by explaining that traditional assessment tools are disadvantageous since high quality tests cannot be constructed easily and the students' productive language skills cannot be checked effectively by these tests. Traditional assessment practices are not enough to decide multidimensional aspects of students' learning skills and they are not effective to assess productive skills and performance of the students. Brown (2004) claims that traditional assessment practices cannot determine several dimensions of students' learning skills and they cannot assess student's performance and productive skills effectively. It just focuses on receptive skills of the learners. Nasab (2015) supports this by saying that students in traditional assessments with selected response do not produce any language. Hence, TA tools are generally used to measure receptive skills of the students. As expressed by Barootchi and Keshavarz (2002), most of the traditional assessment techniques which are not often congruent with present ESL/EFL classroom practices are multiple-choice, norm-referenced, and machine-scorable instruments, if they are used to indicate growth or ability, they may produce faulty results.

True/false, multiple choice, short answer, fill-in-the-blank, matching, and essay are seen as common traditional assessment tools because of their long used in education (Streff, 2016). Stoica (2006) states that in multiple-choice tests students need to choose a response from an already given list. In other words, multiple choice

tests are the most popular traditional assessment tools and commonly used by schools, teachers, and assessment organizations because of the reasons that:

1. “They can be scored fast, easily, and economically. They are actually machine scorable.
2. Objective scoring is applied to them, giving fairer or more reliable test appearance than tests subjectively scored” (Dikli, 2003, p. 13).

Stoica (2006) points out that these tests are rapid and have high reliability of scoring. On the other hand, Simonson et al. (2000) highlight drawbacks of the multiple-choice test tool by stating that it demands low level of cognitive effort. Stoica (2006) indicated the difficulty of constructing multiple choice tests because it is time consuming and hard. In the true-false items, the students should make a choice and find out the true responses (Quansah, 2018). These test items are easy to score and administer. On the other hand, guessing factor increases the chance of finding the correct answer by 50 % when students do not know the correct answer (Cirit, 2014; Quansah, 2018; Dikli, 2003; Simonson et al., 2000). The fill-in and short answer items are written either as direct questions which require the learner to fill in the phrase or a word as statements where a space is left blank for a shortly written answer (Simonson et al., 2000). These types of questions should be precise. If they are not clear, it can be widely interpreted, so students might come up with any possible answers (Simonson et al., 2000). When Stoica (2006) compared multiple-choice and true/false tests with short answer tests, she focused on the benefit of requiring students to give an answer and avoiding cueing not choosing or guessing from fixed options. According to Cirit (2014), matching assessments have a 10% guessing factor which is low but matching can be regarded as a limited method for assessment due to its measurement of the vocabulary knowledge. When essays are taken into consideration, they are accepted very effective in terms of flexibility of the questions and assessment of the higher order learning skills. However, they are not very applicable due to the fact that they are difficult and scoring takes so much time. What's more, subjectivity is significant in scoring, so a rubric can be designed to help teachers grade the essays (Simonson et al., 2000).

2.2.2 Alternative assessment. Wiggins (1990) defines alternative assessment by saying that “assessment is authentic when we directly examine student performance on worthy intellectual tasks” (p. 2). Alternative assessment is a kind of umbrella term

which offers various alternatives to what is commonly used and considered conventional forms of testing. It emerged as a response to the traditional assessment which is not sufficient and specially students and teachers became aware of shortcomings of these tests when they are applied to students who have special goals for learning English. Hamayan (1995) points out that the literature contains alternatives to standardized assessment in several ways: "informal assessment," "authentic assessment," "performance assessment," "alternative assessment," "direct assessment", and "descriptive assessment" (p.213). In this study, the term alternative assessment is used instead of other terms because it has combined characteristics of the other labels which are commonly used (Hamayan, 1995).

Supporters of alternative assessment put it a step further than traditional assessment. According to Nasab (2015), alternative assessment allows the teachers to recognize their students' strengths and weaknesses in changing situations. Authentic assessment is beyond memorization and rote learning. It is much more involving students to perform tasks which include some problem- solving activities (Nasab, 2015). Gökhan (2004) adds that alternative assessment encourage students to use the target language in the same contexts they need to use in real life. Moreover, Brown and Hudson (1998) sum up features of alternative assessment in order to increase awareness of language teachers as follows:

1. ask students to create, produce, perform, or do something;
2. use simulations or real world context;
3. are not intrusive so that they extend the daily classroom activities;
4. allow assessment of students based on their normal daily performance;
5. apply the tasks which represent meaningful instructional activities;
6. focus on products and also processes;
7. understand higher problem-solving and thinking skills;
8. give information about both weaknesses and strengths of students;
9. have multicultural sensitivity when they are properly administered;
10. ensure that not machines but people do the scoring, and human judgment is used;
11. encourage rating criteria and open disclosure of standards;
12. ask teachers to play new assessment and instructional roles.

Quansah (2018) highlights that a student's active generation of a response can be observed either directly or indirectly through a permanent product. Hamayan (1995) adds that for alternative assessment not only process but also product of learning is

important. Nasab (2015) states that alternative assessment give the students better understanding of their achievements and assume more responsibility for their own learning. Barootchi and Keshavarz (2002) give examples of authentic assessment by saying that it includes individual or group performance assessment, observation and portfolios, which can provide a showcase of students' abilities, talents, interests and potentials. With this assessment procedure, some valuable information about students' performance can be provided in educational settings.

There are several concerns about characteristics of alternative assessment. The common concern is related to the validity and reliability of alternative assessment tools. Quansah (2018) states that it carries some concerns based on reliability, subjectivity, and validity. Bailey (1998) gives an example about reliability issue by stating various kinds of students' products may cause reliability concern.

On the other hand, alternative assessment tools depend on the language performance of the students. Gökhan (2004) gives an example on language dependency of alternative assessment tools by saying that a student who is not appropriately proficient in language can hardly perform any tasks. In the alternative assessment, there should be complex thinking skills. Nasab (2015) states that learners also require several guidelines and supervision because of their familiarity with the traditional assessment practices. Nasab (2015) adds another weakness of alternative assessment by stating that it is laborious for the teacher as it takes longer energy and time to decide about the values of the various products presented by the students.

2.2.3 Traditional assessment vs. alternative assessment. Nasab (2015) indicated that assessment diverts its attention from the only use of traditional tests to more authentic assessment methods. That shift is a try to leave static tests and approach more towards real- life tasks. Dikli (2003) states that alternative assessment is used for educational amendment since it is more aware of the effect of testing on instruction and curriculum. On the other hand, Bailey (1998) puts forward that traditional assessment tools are inauthentic and indirect. They are also one shot, norm-referenced as well as speed-based. As Law and Eckes (1995) stated, Dikli (2003) and Quansah (2018) also indicated that traditional assessments are single-occasion tests. It means that they measure performance of learners at a specific time. Supporting what Brown and Hudson (1998) indicated, Herman et al. (1992) also mentioned that alternative assessment methods assess higher-order thinking skills. Gökhan (2004) underlines the

same issue and states that this is an encouraged ability in the current language curricula because students in real life should share their ideas, solve complex problems, and persuade other people, which all needs the use of high-level thinking skills. Winking (1997) claims that authenticity of the alternative assessment tasks encourages the learners to utilize from their higher order thinking skills to solve problems in real life. Dikli (2003) states that students are allowed to show what they learned. The focus of this type of assessment tools is on the student's performance and growth. Development of alternative assessment in context and over time allows the teacher to measure the weaknesses and strengths of the student in different situations. In agreement to Dikli, Hamayan (1995) states that students can see their own achievements and assume responsibility for their own learning when they are assessed by alternative assessment tools. The movement from TA to the alternative assessment was a challenge for the TA methods. The constant changes from the TA to alternative assessment were summarized by Herman et al. (1992) as follows:

- From behavioral to cognitive views of learning and assessment
- From paper-pencil to authentic assessment
- Portfolios: from single occasion assessment to samples over time
- From single attribute to multi-dimensional assessments
- From near exclusive emphasis on individual assessment to group assessment (p.13)

Differences between traditional and alternative assessment are highlighted by Brown (2004) and compared in Table 1.

Table 1

Differences Between Traditional Assessment and Alternative Assessment

Traditional Assessment	Alternative Assessment
One-shot, standardized exams	Continuous long-term assessment
Timed, multiple choice format	Untimed, free response format
Decontextualized test items	Contextualized communicative tasks
Scores suffice for feedback	Individualized feedback and washback
Norm-referenced scores	Criterion referenced scores
Focus on the “right” answer	Open-ended, creative answers
Summative	Formative
Oriented to product	Oriented to process
Non-interactive performance	Interactive performance
Fosters extrinsic motivation	Fosters intrinsic motivation

Traditional and Alternative Assessment adapted from Brown, 2004 pg.13.

It might be easily understood that alternative assessment is superior to traditional assessment.

When it comes to the combination of these two types of assessment, Barootchi and Keshavarz (2002) claimed that standardized measures are widely used for the assessment of students’ progress and should be investigated with other indicators like alternative assessments which show how a language learner learns. Nontraditional assessment or alternative assessment covers anything other than standardized multiple-choice tests. Moreover, traditional tests should be supported by other assessment methods such as alternative assessments which reveal the process of the language learner as learning a language. Due to incongruity of TA methods with the present English language learning practices, alternative ways to assess the students should be searched (Cirit, 2014).

2.2.4 Performance- based assessment. The term “alternative assessment” in the literature is the synonym of other terms such as “performance assessment”, “performance-based assessment”, “authentic assessment”, or “productive assessment” (Berry, 2008, p. 82). Therefore, it is clear that performance- based tasks are an alternative to assessment requiring students to create a product with the purpose of demonstrating their skills and knowledge. It means assessment should involve either a

simulation of a real life activity or observation of a behaviour in the real world (Weigle, 2002).

According to Shohamy (1995), the shift from TA to PBA requires the learners to perform a language which is more authentic and communicative as well as it reflects similar features of real life language use. In addition to that statement, Özuslu (2018) supports this argument by saying that performance- based assessment can encourage students' critical self-evaluation, produce rich learning contexts, increase practitioners' awareness of their students' cognitive processes, simulate real-world problem solving, and cause good instruction by assessing students' abilities to evaluate, solve, and synthesize problems. Overall, we “must rely on observable performances in our assessment of students as much as possible” (Brown, 2004, p. 117). Examining some significant studies in this scope would be beneficial to find out whether PBA methods are effective as an alternative to TA in language classes.

Many scholars conducted some research on PBA. Ozuslu (2018) carried out a study on evaluation of performance– based assessment component of a preparatory class from the perspectives of students, teachers and administrators. Mixed research design was used in order to collect data for the study. When outcomes of the study were taken into account, students had positive attitude towards the PBA while teachers and administrators had more positive towards PBA. Additively, all participants of the study agreed on the continuity of the PTs in the next years.

Chen (2007), from University of New Orleans, conducted a Ph. D. dissertation with college students on the purpose of investigating learners perceptions on differentiated instruction offering tiered PTs. Although the study provided some positive outcomes such as improved motivation, greater confidence, increased efforts, some negative findings were also obtained such as limited time, scoring and difficult tasks. Additively, Örsdemir (2010) carried out a study with state primary school students in order to find out how PTs were implemented in EFL classes. Questionnaires, analysis of documents and interviews were used as data collection instruments. Findings of the study revealed that PTs were beneficial practices focused on increasing motivation, improving vocabulary, and increasing consciousness of the learners.

What's more, Gökhan (2004) investigated teachers' and students' opinions about their involvement in the assessment process and about the use of alternative forms of assessment. The study was conducted with 50 teachers and 120 students at Hacettepe

University Basic English Division. The findings of the study revealed that teachers and students find the assessment system satisfactory in itself; however, the present system has weaknesses as well as strengths. This emphasizes the importance of supplementing the weak points with alternative forms of assessment and using multiple assessment methods to obtain more effective results.

Another recent study conducted by Koné (2015) examined the impact of PBA on university ESL learners' motivation. Additively, the researcher investigated the difference between students who had more experience and who had less experience with performing an oral presentation group project. The results of the study revealed that the students responded positively to that type of project. However, their motivational and emotional states varied across time depending on their experience with the oral presentation, their performance, and the cohesion of their group.

Bas (2011), moreover, examined the effects of project-based learning on ninth grade students' academic achievements and attitude towards English lessons in EFL context. To reach this objective, data were collected from 60 Turkish English learners when they were dealing with a unit called past activities. The participants' academic achievement test related to the English curriculum and the English lesson attitude scale were used to elicit data from these respondents. English lesson attitude scale measures these participants' attitude towards English lessons. The results indicated that there was a significant statistical difference between the experimental group and control group achievement. Concerning their attitude towards English lessons, the findings revealed that the experimental group had a more positive attitude and higher performance than the control group.

Islam and Ahmed (2018) explored the current status of the practice of alternative assessment and its effectiveness in tertiary level ESL classroom in Bangladesh by conducting a research among the participating students and teachers. Additively, the researchers investigated the effectiveness of alternative assessment depend on the level of the learners. They utilized from both qualitative and quantitative research designs for collecting and analyzing the data.

When all studies mentioned above are taken into the consideration, the results show that students mostly have positive attitude towards performance-based assessment tools.

Chapter 3

Methodology

This chapter aims to describe methodology of the study by providing information related to research design, setting, participants, procedures, and limitations and delimitations of the study.

This study aims to answer these research questions:

1. What are the students' perceptions of the PTs regarding task familiarity, application, scoring, learning outcomes, and program consistency?
2. What are the impacts of PBA tasks on students?
3. Do the students' perceptions of PBA tasks vary based on their proficiency level?

3.1 Research Design

The goal of this study is to search for the perceptions of students towards performance-based assessment as an alternative assessment. As clearly seen from the literature, most researchers who studied on comparison or combination of assessment tools generally utilized mixed method research design. It is mostly defined as combination of quantitative and qualitative research designs. In the same vein, Johnson and Onwuegbuzi (2004) point out that the main purpose of the mixed methods is to minimize the weaknesses and draw from strengths of the studies. Within the scope of this study, the researcher utilized both quantitative and qualitative research design with the aim of obtaining sufficient data.

Quantitative data of the research was obtained through a 5 point Likert- scale. Dörnyei (2003) states that quantitative research uses viewpoints, categories, and models as the essential characteristics which are precisely defined by the researcher in advance. Also, quantifiable or numerical data are collected to test the research hypotheses. A broader perspective has been adopted by Çelik (2015) who claims that quantitative findings in social sciences can be regarded as unsatisfactory and problematic and should be further reevaluated. Additively, the qualitative findings allow the researcher to make comment on the descriptive findings.

The data obtained from focus discussion group practice were analyzed in order to get qualitative data. Qualitative research is defined as an approach:

“ for exploring and understanding the meaning individuals or groups ascribe to a social or human problem. The process of research involves emerging questions and procedures, data typically collected in the participant’s setting, data analysis inductively building from particulars to general themes, and the researcher making interpretations of the meaning of the data. Those who engage in this form of inquiry support a way of looking at research that honors an inductive style, a focus on individual meaning, and the importance of rendering the complexity of a situation” (Creswell, 2013, p. 68).

In order to gain better insight while analyzing qualitative data, a thematic analysis method was used. Maykut and Morehouse (1994) defined it as a kind of method aims to analyze qualitative data by categorizing and coding with the same units of meaning.

By combining both quantitative and qualitative research methods, the researcher aimed to get the validated findings which were reviewed in great detail to answer the research questions.

3.2 Setting and Participants

3.2.1 Setting. This study was conducted with secondary level ESP students at a private vocational and technical anatolian high school in İzmir during the Fall Term of the 2019- 2020 academic year. The participant students were all 11th grade. The institution where the study was conducted is one of the private vocational and technical high schools which has two programs; aircraft maintenance technician program and aviation logistics program. This study focused on students who study at aircraft maintenance technician (AMT) program. When the students graduate from this school, they get a chance to study at some aircraft programs offered by some distinguished universities in Turkey. On the other hand, if they don’t want to study at university after graduating from high school, they may get a job with the certificate provided by the school. Basically, an aircraft maintenance technician is responsible for maintenance, repair and installation of aircraft engines and systems. They have to follow Aircraft Maintenance Manual (AMM). It gives the procedures in a detailed way in order to carry out the inspection, maintenance and repair of aircraft engines, electronic and

avionics systems. When the AMT program students are 9th grade, they take general courses like mathematics, literature and English before they go to their departments. Their departments are determined according to the students' total score of 9th grade courses. When they achieve 70 points out of 100, they get a chance to study at the AMT program. The students' class level is also determined based on the average of the grades they get during 9th grade. During 3 years of the program, they all take technical classes such as basic aerodynamics, electronic device systems and human factors. In line with these courses, they have vocational English class which were offered just 2 hours per week.

The data for this study was gathered from 11th grade students taking the course vocational English throughout the whole academic year. This study includes 52 students from three different classes. The researcher was also the vocational English instructor of those three groups. All classes followed the B1 level course book which was provided by the MoNE. On the other hand, some students were quite below the assumed level of language proficiency as the administration of the school formed classes that were used in the study according to their scores.

The assessment type applied in the school was generally through the application of two standardized exams (50%), one term paper (25 %) and one classroom performance grade (25 %) in each semester. Performance tasks were decided and designed depending upon the students' needs which were observed by the researcher previous year. Only doing the four tasks designed for the study was obligatory since the grades the students get from the tasks will impact 25 % of their overall grade in the course. The teacher assigned one term paper to those who did not want to participate in the study.

Before collecting the data, consent form of the study was given to the students by stating that signing the consent form or filling the survey was not obligatory, so they didn't have to fill the form. The students were asked not to put their names on the form so as to make them feel honestly while taking the questionnaire. The volunteers filled the form, completed the questionnaire and submitted it to the researcher.

The reason why this study was conducted at that school was because the researcher was teaching full time at that institution. Hence, conducting the research at that institution was the only possibility because of the researcher's convenience. However, it was an advantage for students because when they came across a problem

while dealing with the tasks, they had a chance to consult the researcher so as to clarify the issue and come up with a solution to the problem.

3.2.2 Participants. The participants of the study were 8 female and 44 male 11th grade students studying in a private vocational and technical high school. There were three different language proficiency levels, namely elementary, pre-intermediate and intermediate as the classes were formed according to students' academic scores. Total number of 11th grade students was 73 because 21 of them did not want to participate in the study or did not fulfill all the requirements of the study. All participants were native speakers of Turkish and graduated from state or private primary schools where they studied English for 4- 8 hours per week. The students' age ranged from 15 to 16 years. All participants in the study were studying in their second year at the vocational English course. Because of the confidentiality issue, the identities of the participants were kept secret.

The frequency and percentage values of the demographic data of the students in the sample of the study are shown in Table 2.

Table 2

Demographic Data of the Participants of the Study

Gender	F	%
Male	8	15,4
Female	44	84,6
Total	52	100,0
English Proficiency Level		
Elementary	10	19,2
Pre-Intermediate	22	42,3
Intermediate	20	38,5
Total	52	100,0
Performance- Based Assessment Experience with the Vocational Program		
Yes	21	40,4
No	31	59,6
Total	52	100,0

Table 2 shows that there are 52 students in the sample. 85% of the students are male students and 15% of the students are female students. When the students' current English Proficiency level in the study is taken into account, it is seen that 19% are elementary, 42% are pre- intermediate and 38% are intermediate. As students' classes are determined on the basis of their proficiency level, there are three levels in this study. The demographic data shows that 40% of the students in the study were previously evaluated through one of the PBA methods in the vocational English classes, but 60% were not.

3.3 Procedures

This chapter provides information about data collection instruments, data collection procedures, data analysis procedures, trustworthiness, validity, and reliability of the present study. At the end of the chapter, limitations and delimitations of the study are presented and discussed.

3.3.1 Data collection instruments. Two different tools were used in order to collect data for this study. When all students completed all performance based tasks, a questionnaire was carried out in order to learn students' attitude towards the implemented performance based tasks' task familiarity, application, scoring, learning outcomes and program consistency sections. At the end of the term, a focus group discussion was formed with 10 students who were randomly selected.

3.3.1.1 Questionnaire. This data collection instrument included two components as well as a consent form. In the first part of the questionnaire (Part A), the participants' demographic data was collected to learn their background. According to Brown and Rodgers (2002), Likert scales are generally utilized in order to get information about participants' attitude towards various aspects of language learning. Hence, in the second part (Part B), a 5 point Likert- scale was used in order to investigate participants' perceptions on PT implementation. The scoring of the questionnaire can be explained as follows: 5 = Strongly agree, 4 = Agree, 3= Undecided, 2 = Disagree and 1 = Strongly disagree. The questionnaire had five constructs which were task familiarity, application, scoring, learning outcomes, and consistency with the vocational English program. In task familiarity phase of the implementation, the

researcher aimed to find out students' attitude towards the PTs, scoring rubrics and investigated readiness of the students for the tasks before implementing PTs. In the next construct, students gave some detailed information about "application" process and challenges of the PTs. "Scoring" construct had to take part in this process because it was an assessment. The purpose of "learning outcomes" construct was to investigate for contribution of PTs to the students. In the last construct, students' perceptions on the consistency of the PTs with the content of the vocational English course was aimed to be learned through the use of the questionnaire.

The questionnaire mentioned above was developed by Ozuslu (2018). During literature review, the researcher came across it which was an appropriate survey fulfilling the needs of the present study in terms of data collection instruments. It was used in the present study after taking consent of the writer and writer's advisor. The researcher of present study made necessary changes to the survey regarding the feedback given by the advisor of the study. For example, as a title for the items from 1 to 8, the researcher used the word 'task familiarity' instead of 'planning' in order to clarify the aim of the construct.

The developer of the questionnaire investigated the psychometric values of the measurement tool. The reliability of the quantitative was tested in SPSS by using Cronbach's Alpha. Cronbach's Alpha value of the the entire questionnaire was rated at 0.93, which indicated that the questionnaire had a high level of reliability. Kaiser-Mayer-Olkin (KMO) and Bartlett's test (BTS) were carried out for the validity check of the questionnaire. As KMO value was high (, 838) and BTS result of the scale showed that p value was (, 000) lower than, 01, the scale was suitable for factor analysis. With the aim of checking the construct validity of the scale, factor analysis, principle components analysis, was administered. The total variance of the five factors whose attributes were higher than 1, 00 was % 53,150 for the scale. It was a valid analysis according to Ozuslu (2018) who explained that "an analysis which explains %50-75 of the total variance was a valid analysis but %40 and higher can also be considered as a valid" (p. 74).

3.3.1.2 Focus group discussion with students. Focus groups are used as one of data collection tools in qualitative research. According to Ary et al. (1996), participants of the study will be free to express their opinions in their own words and comment on the responses of the researcher and other participants while conducting focus group

discussions. The interaction between the students helps them to generate ideas and express their opinions freely, so the researcher gets more valuable data about the subject matter (Kesmen, 2018).

Within the scope of this study, focus group discussion was the second instrument used to gather data for the study and was conducted with 10 randomly selected students one week after the survey at the end of the term. It lasted approximately 30 minutes. It was conducted in the school environment and audio- taped by the researcher.

3.3.2 Data collection procedures. The data gathered for the aims of this study was analyzed both quantitatively and qualitatively. With the consent of school administration, the study was conducted in vocational English course by the researcher who was the teacher of the lesson. The researcher prepared all PTs along the same line with schedule and course content. This study was conducted in the Fall Term of 2019-2020.

3.3.2.1 PT implementation process. While designing the tasks, the researcher puts emphasis on the consistency between the tasks and content of the vocational English course. PTs used in this study were chosen on the basis of the course requirements and students' needs which was observed and investigated by the researcher in the previous years. It was observed that students could not transfer their structural or lexical knowledge into written and oral output. For example, the researcher designed listening- speaking, writing and vocabulary tasks as they were predominantly practised in vocational English class. Through the PBA tasks, in accordance with the objectives of the program, it was aimed to enhance their productive skills plus the receptive ones. PBA tasks, by their nature, were formed to foster all language skills in an integrated way. To exemplify it within the scope of vocational English PTs, for speaking task, students were expected to read the course book and Aircraft Maintenance Manual (AMM), write down the important points regarding their topic, prepare a PPP, present it in the classroom, and listen to the presentations given by other students. Four language skills were activated and aimed to improve by this way. Motivation, autonomy and 'higher-order thinking skills' of the students were aimed to be enhanced by the use of motivating tasks.

As there were four PTs in this study, the researcher did not want to implement a task every week not to put so much pressure on the students and also to give the

necessary time for the completion of each task. Hence, at least 2 weeks were allocated for each task. Weekly schedule of the PTs in vocational English course was given in Table 3 below.

Table 3

Weekly schedule of the PT implementation

Week	Date	PT
1	September 9, 2019	-
2	September 16, 2019	Introduction week Assign task 1
3	September 23, 2019	-
4	September 30, 2019	Task 1
5	October 7, 2019	Assign task 2
6	October 14, 2019	-
7	October 21, 2019	Task 2
8	October 28, 2019	Assign task 3
9	November 4, 2019	-
10	November 11, 2019	Task 3
11	November 18, 2019	Fall break
12	November 25, 2019	-
13	December 2, 2019	Assign task 4
14	December 9, 2019	-
15	December 16, 2019	Task 4
16	December 23, 2019	Questionnaire
17	December 30, 2019	Focus group discussion
18	January 6, 2020	-

Introduction Week: (Sept 16)

- Brief information was given to the students about the task implementation process.
- Students were informed about the weekly schedule of the tasks.
- Importance of the scoring rubrics and their usage was explained in a detailed way.
- Grading system of performance based tasks in vocational English course explained to the students. Each PT was calculated over 25 points.

Task 1: (Sep 30)

- It was a role- play task. Students covered basic hand tools unit in the previous weeks. As this task was a pair work, one of the students became an aircraft maintenance technician and the other one became an intern in an aircraft company. AMT asked some questions to the intern regarding the usage of same basic hand tools. AMTs were required to make detailed explanations if the intern did not provide a reasonable response. Hence, both students were required to master the subject.
- Students worked in pairs.
- Time limitation for each group was 5 minutes.
- It was calculated over 25 points.
- Feedback was given by the researcher right after the implementation.
- Speaking scoring rubric was used by the teacher (Appendix E).

Task 2: (Oct 21)

- It was a writing task. A video used in this task showing that a component of an aircraft did not work properly. In other words, some technical problems occurred in the component. Hence, students wrote a report about it on the basis of problems and provided possible solutions with the guidance of AMM.
- Each student wrote a paragraph between 100 and 150 words.
- If a student hands the report in one day later than instructed, their grade drops by 10 %.

- The teacher shared writing rubric (Appendix C) with the students while explaining the details of the task.
- After assessing students on the basis of writing rubric, feedback was also provided to the students by the researcher.

Task 3: (Nov 11)

- Students were free to pick one of these topics related to parts of an aircraft which was covered in the class and present it by designing a Power Point presentations (PPP) or Prezi presentation.
- Individual work was required in this task.
- Each student had just 5 minutes to present their subject.
- It was calculated over 25 points.
- Speaking scoring rubric (Appendix E) was used by the researcher.

Task 4: (Dec 16)

- As the last task was performed at the end of the term, the students needed to revise all vocabulary items they had covered throughout the semester, so the researcher asked the students to prepare a quiz via Quizlet to which students got accustomed to use in the class. (5 points out of 25)
- After presenting the required vocabulary items via Quizlet, each student prepared a quiz show “who wants to be a millionaire?” to test given vocabulary items (20 points).
- Students chose 20 words in order to present from the relevant sections of the book. However, the students prepared a quiz show in which he/she asks 10 questions. The target words could be involved in either the questions or in the answers.
- Each student had 15 minutes.
- Vocabulary scoring rubric was used to assess students’ performance (Appendix D).
- Detailed feedback was given to students on the basis of the activity they designed.
- It was calculated over 25 points.

3.3.2.2 Grading and scoring rubrics. According to the grading system of MoNE, application of two standardized exams, one term paper and a classroom performance grade is compulsory in vocational English class in each semester. Each of them was calculated over 25 points. Grading system of vocational English class was illustrated in Figure 2 below.

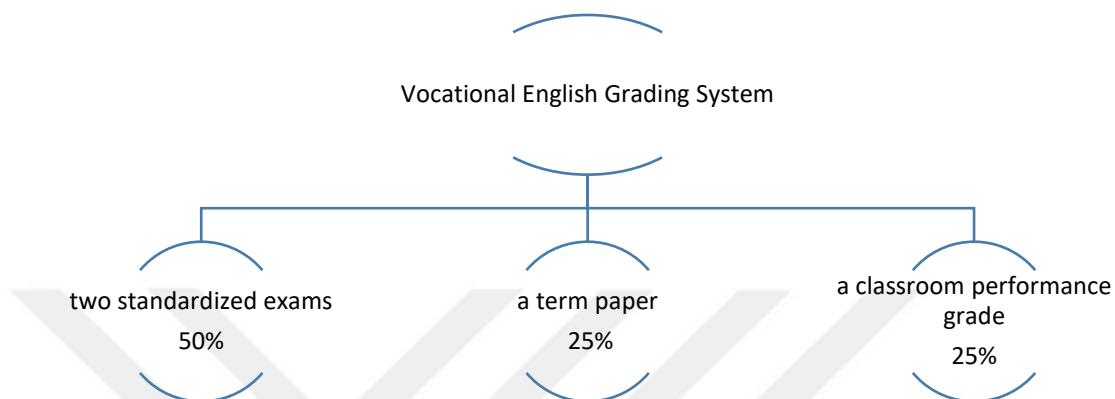


Figure 2. Grading system of vocational English class

At the very beginning of the term, the researcher explained all details about the grading system in order to clarify whole process for the participants. There were 4 tasks in the performance based assessment system. Each of them was assigned with 25 points. Total score of the tasks accepted as a term paper grade of the students, which composed 25 % of the general score average. Each task was calculated out of 25 points. Then overall average of the tasks divided into 4 in order get final point of the student.

The researcher made the use of different scoring rubrics as the PTs are different from each other and shared them with the students with the aim of giving information about the requirements of the PTs. The researcher used reliable scoring rubrics in order to grade PTs of the students. Scoring rubric for writing was designed by the teachers in the institution in the previous years. Speaking and vocabulary PTs were assessed through the scoring rubrics generated by Ozuslu (2018). They were calculated over a hundred.

3.3.3 Data analysis procedures. As a quantitative data collection tool, the researcher used a questionnaire so as to find out participants' attitude toward the performance- based assessment. As suggested by Brown and Rodgers (2002), any research questions requiring description, explanation, or exploration of people's

attitudes, characteristics, views, and opinions can be answered by using the language surveys.

With the aim of analyzing quantitative data, IBM SPSS was used in order to measure descriptive statistics including means, frequency, standard deviation, and percentages. The quantitative data which was collected through the 5 point Likert-scale which has 45 items including 5 phases such as task familiarity, application, scoring, learning outcomes and program consistency. The scoring can be explained as follows: 5 = Strongly agree, 4 = Agree, 3= Undecided, 2 = Disagree and 1 = Strongly disagree.

For the qualitative part of the research, thematic analysis was used in order to evaluate PBA and understand attitudes of students toward the performance- based assessment in vocational English courses. As a way of summing up parts of the data, thematic analysis was defined by Saldaña (2016) as a way through which those summaries are grouped into a smaller number of themes, categories, or concepts. The data was collected in Turkish because the researcher aimed to reach more detailed data. The answers of the students were translated into English with an expert in Translation and Interpreting. The researcher analyzed all qualitative data and coded the answers of the students with the aim of finding meaningful themes. To this end, these codes were categorized according to suitable themes.

3.3.4 Validity and reliability. For reliability of the quantitative research, Crocker and Algina (1986) state that test developers should demonstrate whether the scores of their tests are reliable or not. For the purpose of enhancing reliability, the researcher made sure that all items were easy to understand and clear for the participants of the study. While conducting the survey, participants were allowed to ask their questions to the researcher if they felt confused. The reliability of the quantitative data which was gathered through the survey was tested in SPSS by using Cronbach's Alpha. Table 4 illustrated detailed test results for Cronbach's Alpha.

Table 4

Cronbach's Alpha

Data	
Cronbach's Alpha	,95
N of Items	45

At the end of the implementation, the reliability of the questionnaire was tested by Cronbach's alpha using SPSS and α value was measured 0.95. The criteria for the reliability of the questionnaire is given by (Ozdamar, 2004, p.633) in Table 5 below.

Table 5

Reliability criteria for α value

α value	Reliability of the tool
$0.00 \leq \alpha < 0.40$	No reliability
$0.40 \leq \alpha < 0.60$	Low reliability
$0.60 \leq \alpha < 0.80$	Quite reliability
$0.80 \leq \alpha < 1.00$	High reliability

Reliability score of the scale indicates a high and acceptable level of reliability for the questionnaire.

3.3.5 Trustworthiness. Trustworthiness is very crucial for qualitative research. Seale (1999) puts emphasis on the significance of validity and reliability in qualitative research in order to carry out strong studies by stating that the research report shows trustworthiness for the issues conventionally discussed as reliability and validity. According to Lincoln and Guba (1985), in establishing trustworthiness there are four main components which are credibility, dependability, confirmability, and transferability.

Credibility is one of the significant criteria established by positivist researchers regarding internal validity of a qualitative research. It aims to ensure whether the study measures what is actually intended. In other words, it is all about the truth of the findings of a study. Triangulation was a crucial method used in order to increase

credibility of present research. The data was collected through different instruments so as to get more detailed data.

Transferability is a form of establishing external validity of a research. It is established by applying the results of a study to other contexts and wider populations. This research can be a sample for future researchers as there is limited research which provides details on PTs used in ESP classes and students' attitude towards PTs as a way of alternative testing.

Dependability discusses if this study was repeated with the same participants, context and methods, similar results of the study would be obtained. In other words, it concerns consistency of the findings. In this study, different data collection instruments and detailed data analysis were used with the aim of revealing the consistency of the findings.

Last strategy to ensure trustworthiness in a research is confirmability which ensures that the work findings are obtained from the ideas and experiences of the informants, not from the researcher's preferences and characteristics (Shenton, 2004). That is to say, it concerns to the degree which the study findings could be confirmed and corroborated by others. One of the essential ways of enhancing confirmability, which was established in this research, is triangulation. It helped to lower the researcher bias. Also, another one is member checking. All transcribed data were also analyzed with one of the colleagues of the researcher from the field of ESP with the aim of establishing inter-coder reliability. In order to enhance reliability of the study, the outside coder and the researcher compared their codes and agreed on them.

3.4 Limitations and Delimitations

This study has limitations and delimitations. One of the main limitations of this study is that the number of the participants was 52 but the number could have been higher. In this way, the sample can be generalized to a larger group and its scope can be further widened in the future, so reliability of the research could be improved. Another limitation occurred during focus group discussion with participants from different classes. Because of the reason that some students were eager to speak first and most, some students preferred being quiet instead of expressing their own opinions.

As a delimitation, it can be true to say that the study is one of the first ones implemented on vocational high school students' attitudes toward performance based assessment as an alternative assessment.



Chapter 4

Findings

A questionnaire and focus group discussion methods were implemented in order to investigate students' opinions concerning PBA. The questionnaire involved 45 Likert- type items. These items were categorized into 5 sections which were task familiarity, application, scoring, learning outcomes and program consistency.

Findings of the questionnaire and focus group discussion aim to reach answers of the following research questions:

1. What are the students' perceptions of the PTs regarding task familiarity, application, scoring, learning outcomes, and program consistency?
2. What are the impacts of PBA tasks on students?
3. Do the students' perceptions of PBA tasks vary based on their proficiency level?

4.1 Findings of the Research Question 1: What are the students' perceptions of the PTs regarding task familiarity, application, scoring, learning outcomes, and program consistency?

For the purpose of answering the last sub-question of the research, a questionnaire with 5 point Likert- type items was implemented to 52 students. Students' overall opinions about the five sections of the questionnaire and number of items for each phase of the study are presented in Table 6 below.

Table 6

Descriptive Statistics Results of Students' Ideas Related to Phases of the Performance-based Assessment.

	NI	\bar{X}	Sd
Task familiarity	8	4,25	,534
Application	10	3,88	,504
Scoring	8	4,12	,559
Learning Outcomes	12	4,15	,617

Table 6 (cont.d)

	NI	\bar{X}	Sd
Program Consistency	7	3,65	,463
Total	45	4,03	,518

Table 6 shows the mean values of the students' ideas regarding the performance tasks implemented at school. It is clearly seen that students generally evaluated the performance based assessment positively with an overall mean 4.03. It was observed that the highest mean score of performance implementation stages was in task familiarity with 4.25. That was followed by learning outcomes phase ($\bar{X}= 4.15$), scoring phase ($\bar{X}= 4.12$), application phase ($\bar{X}= 3.88$) and the lowest score was consistency with the program ($\bar{X}= 3.65$) successively. In general, it was seen that students had a positive attitude towards the evaluation of performance based assessment.

In order to learn about students' perceptions towards task familiarity section, the items 1 through 8 were analyzed in detail. Table 7 illustrates the findings of the analysis.

Table 7

Perceptions of the Students on the Items Related to Task Familiarity Phase

Items	N	\bar{X}	Sd
1. Instructor had sufficient information about PT implementation.	52	4,85	,607
2. We were sufficiently trained before the PT implementation.	52	4,54	,726
3. Instructor sufficiently knows the scoring rubrics.	52	4,79	,667
4. I did not know how to prepare PTs.*	52	1,83	,879
5. We are given clear instructions about each PT by the instructor.	52	4,52	,851
6. There was a weekly schedule for each PT.	52	4,83	,648
7. We were notified of the deadlines to give PTs.	52	4,85	,638
8. I as a student felt to be ready for PTs before the implementation of PT.	52	3,85	,894
Total	52	4,25	,534

Opinions of the students regarding the items in the task familiarity stage of the performance implementation are shown in Table 7. Students expressed that they were pleased with the task familiarity phase of the implementation ($\bar{X}=4,25$). Considering the task familiarity items, it was seen that the teacher had sufficient knowledge about performance practices and made sufficient explanations about performance task implementation to the students. The low mean score of item 4 * ($\bar{X}=1,83$) illustrated that students understood how to prepare performance tasks. To put it briefly, students had positive attitude towards the task familiarity phase of the implementation.

In the second part of the questionnaire, the purpose of the researcher was to find students' attitude towards the application phase of the PTs. Table 8 shows the analysis regarding all the answers to the questions from 9 to 18.

Table 8

Perceptions of the Students on the Items Related to Application Phase

Items	N	\bar{X}	Sd
9. PTs fitted with my language level.	52	4,42	,825
10. PTs fitted with my cognitive level.	52	4,27	,888
11. We enjoyed PTs.	52	4,65	,738
12. We were given sufficient time to complete PTs.	52	3,90	,748
13. It was time-consuming to implement PT.	52	1,58	,915
14. We were sufficiently guided during PT process.	52	4,73	,843
15. I was bored of doing PTs.*	52	1,77	,983
16. I saw presenting PTs in class as effective.	52	4,13	,908
17. I could give my PTs immediately.	52	4,63	,817
18. I myself prepared my PTs.	52	4,71	,800
Total	52	3,88	,504

**Reverse Item*

Opinions of the students related to the items in the application phase of the PT implementation are indicated in Table 8. Students clearly stated that they had highly positive attitude towards the application phase of the study ($\bar{X}=3,88$). The highest score was seen in the 14th item with the highest mean of ($\bar{X}=4,73$), so it could be stated that sufficient guidance was provided during the performance based assessment.

According to the mean score of the 18th item ($\bar{X}=4,71$), students prepared the performance tasks on their own. Moreover, most of the students agreed with the 11th item ($\bar{X}=4,65$) and 17th item ($\bar{X}=4,63$) which revealed performance tasks were fun and they managed to submit performance tasks on time. The low mean value of item 13th * ($\bar{X}=1,58$) and 15th item * ($\bar{X}=1,77$) revealed that according to the students performance tasks were not waste of time and they were not bored while preparing them.

In the third part of the questionnaire, the perceptions of the participants towards scoring phase of performance- based assessment tasks were revealed through items from 19 to 26 in Table 9.

Table 9

Perceptions of the Students on the Items Related to Scoring Phase

Items	N	\bar{X}	Sd
19. A rubric was used to evaluate each PT.	52	4,75	,837
20. Scoring rubrics could be easily understood.	52	4,25	,883
21. PTs were scored objectively.	52	4,67	,810
22. PT scores were satisfactory.	52	4,10	,774
23. There was sufficient time for assessment.	52	4,08	,763
24. PTs were not scored fairly.*	52	1,44	,777
25. We were given feedback about our PTs.	52	4,81	,817
26. There was planned scoring of both product and process of PTs.		4,85	,638
Total	52	4,12	,559

**Reverse Item*

Students expressed that they were satisfied with the scoring phase of the performance- based assessment ($\bar{X}=4,12$). The highest mean value of the phase was item 26 ($\bar{X}=4,85$) which showed that students agreed with the idea that there was planned scoring of both product and process of PTs. When the order of the mean scores were taken into account, the second highest value was 25th item ($\bar{X}=4,81$) which revealed that students got enough feedback about their performance tasks by the instructor. Moreover, item 19 ($\bar{X}=4,75$) and item 21 ($\bar{X}=4,67$) shed light the fact that each PTs were evaluated through a rubric and they were scored in an objective way

successively. The lowest mean value of the scoring phase was item 24 ($\bar{X}=1,44$) which means that according to the students PTs were scored fairly.

In the next part of the questionnaire, students' attitude towards the learning outcomes phase of the PTs were analyzed. In Table 10, the analysis of the results shows that all the answers to the questions from 27 to 32.

Table 10

Perceptions of the Students on the Items Related to Learning Outcomes Phase

Items	N	\bar{X}	Sd
27. PTs contributed to developing my ESP writing skills.	52	3,42	,871
28. PTs contributed to developing my ESP reading skills.	52	4,18	,745
29. PTs contributed to developing my ESP speaking skills.	52	4,06	,752
30. PTs contributed to developing my ESP listening skills.	52	3,25	,764
31. PTs contributed to developing my ESP grammatical knowledge.	52	4,37	,817
32. PTs contributed to developing my ESP vocabulary knowledge.	52	4,81	,658
33. PTs contributed to developing my general language proficiency.	52	3,40	,721
34. PTs increased my autonomy.	52	4,48	,852
35. PTs increased my motivation to learn English.	52	4,62	,820
36. PTs increased my self-confidence.	52	4,27	,952
37. I think that PTs did not contribute to my general language proficiency.*	52	1,52	,779
38. I reinforced what I have learned in course through PTs.	52	4,29	,776
Total	52	4,15	,617

**Reverse Item*

When it comes to the items in learning outcome phase, the opinions of the students are shown in Table 10. According to that table, general attitude of the students was mostly positive about learning outcome stage of the study with the mean of ($\bar{X}=4,15$). The highest mean score of item 32 ($\bar{X}=4,81$) showed that performance based assessment contributed to the development of ESP vocabulary knowledge of the students. It was followed by 35th item ($\bar{X}=4,62$) which indicated that performance

based assessment increased students' motivation to learn ESP and item 34 ($\bar{X}=4,48$) revealed that students learning autonomy increased with the application of performance based assessment. Students all agreed and had positive attitude towards contribution of the performance based assessment to the development of reading, speaking and writing skills. It was revealed that they were quite not sure about PTs' contribution to their listening skills Students disagreement to the 37th item * ($\bar{X}=1,52$) revealed the fact that performance tasks contributed to their general language proficiency.

In the last phase of the questionnaire, the attitudes of the students towards the PTs were investigated. Table 11 illustrates the analysis of the items from 39 to 45.

Table 11

Perceptions of the Students on the Items Related to Program Consistency Phase

Items	N	\bar{X}	Sd
39. Performance-based assessment is an effective assessment method for ESP.	52	4,40	,774
40. Performance- based assessment methods are more helpful than traditional ones for language teaching and learning.	52	4,14	,727
41. There was consistency between the PT implementation in this institution and the course content.	52	4,83	,648
42. Performance-based assessment elements followed the objectives of the Technical English program in this school.	52	4,83	,734
43. Implementation of PT in this institution showed the principles of performance-based assessment.	52	4,19	,886
44. Implementation of PT in this institution did not show our learning.*	52	1,13	,627
45. Performance- based assessment is useful only when they supplement the traditional assessment.	52	3,63	,768
Total	52	3,65	,463

***Reverse Item**

Generally, Table 11 shows that general attitude of the students toward the items regarding program consistency was mostly positive ($\bar{X}=3,65$). Item 41 ($\bar{X}=4,83$) got the highest mean value which shows that implementation of PT in this school matches

with the course content. Additionally, students stated that elements of performance-based assessment followed the objectives of the vocational English program of this school in the 42nd item with the value of ($\bar{X}=4,83$). Students were not agree with the 44th item * ($\bar{X}=1,13$) and this low mean score indicated that PT implementation reflected what students have learned in vocational English course. Last item revealed the fact that students preferred performance-based assessment as a supplementary tool to traditional assessment.

4.2 Findings of the Research Question 2: What are the impacts of PBA tasks on students?

In this chapter of the study, the researcher collected qualitative data through forming focus discussion group with 10 students. Some essential themes appeared when the data were analyzed by the instructor. According to the data obtained from focus group discussion, some benefits and challenges of the performance based assessment emerged. While analyzing the data of the focus discussion group, the text was divided into segments of information of codes for development of a more general image of the data. At the beginning, twenty-five codes were found as open coding. Later, in the axial coding phase, it was reduced to eight, and finally in the selective coding phase, the codes were reduced to two themes as benefits and challenges of PTs'. In addition to this, sample interview excerpts of some students were included in the paper.

4.2.1 Benefits of using PBA tasks. In the light of data obtained from data collection instruments, benefits of being assessed through PTs as an alternative way of traditional assessment emerged in 3 sub- themes;

1. Increasing motivation of the students
2. Fostering language development
3. Developing awareness and social skills of the students

4.2.1.1 Increasing motivation of the students. When the responses of the students taken into the consideration, it is seen that motivation of the students stimulated by the PTs and also students' attention triggered by the tasks. Their attitudes and behaviors became more attentive and positive towards PTs as they were getting

used to the performance based assessment. It was concluded that students ended up using the target language more efficiently in the classroom environment and being more enthusiastic to prepare and present their tasks. As Kesmen (2018) stated “intrinsic motivation is a key element to keep on supporting motivation and finding situations and ways to foster it” (p.46). As seen in the examples below, students realize their effort and the necessity of maintaining it on their own. PBA tasks are more concrete, it helps students gain a conscious awareness by giving them the opportunity to apply them in their daily lives which also contributes to the increase in their motivation. Having the opportunity to transfer what they have learned into their daily lives, their motivation increases. As implied from the remarks of the participants, being motivated was the key to keep the enthusiasm alive. Responses of the students about motivation are given below:

“I liked tasks because I had fun while preparing and presenting them” (Student 5)

“I felt more motivated and self-confident while I was preparing tasks especially role- playing task.” (Student 3)

“When the teacher first introduced the performance based assessment, I was scared but then I got used to it and felt more motivated to prepare them. I did a lot of research on the net to prepare better tasks.”(Student 6)

“I started to spend more time on projects and lessons. I am studying more. I am motivated to learn more.” (Student 1)

4.2.1.2 Fostering language development. Contrary to the traditional system, through PBA, students have given an opportunity to be at the center of their learning process. According to the responses of the students, students feel under pressure with the traditional exams because they have limited time. However, when PTs were taken into account, their attitude was more positive because they got a chance to express themselves better and they had enough time to prepare each task. They also indicated that PTs contributed their learning. Additionally, remarks of the students indicated that PTs improved their higher order thinking skills.

“ We searched for the subject on the internet and improved the lecture notes to perform the task. The lesson content became permanent in this way and helped us become successful. Otherwise, the subjects would not be revised by a student after the class.” (Student 5)

“ The course became more enjoyable by performing the tasks. It made us revise our understanding of the subject”(Student 2).

“ Most of our knowledge could be reflected easily through PTs. They were better than we do in the one hour traditional exams”(Student 9)

“ Even though we give very much effort while preparing tasks, it is very fruitful to have a chance to apply what we have learned.” (Student 7)

When the answers of the students analyzed, it is clearly seen that PTs helped them to learn from each other. Tasks keep the students updated in terms of the course content. Therefore, they are more actively involved to the lessons. On the other hand, students stated that it was a good way of learning from each since they had a chance to observe their friends' presentations in class. Comments of the students about this issue are given below:

“ During my friends' presentation, I have learned so many new things regarding the subject they present.” (Student 1)

“ Especially practicing new vocabulary items with vocabulary games and using Quizlet helped me learn new words” (Student 10)

“ I have learned many things from my friends' presentations. However, the main thing I have learned was how to do presentation in front of other people.” (Student 7)

When remarks of the students were taken into the consideration, it would not be too wrong to say that PTs helped students to practice ESP language in an authentic way. As the students were assessed by using authentic tasks which gave them an opportunity to enhance authentic language. The participants pointed out that with alternative assessment, authentic materials could be used and students' language skills could be increased. Students were glad to perform the job specific tasks since they were aware of the fact that those skills attained by the tasks would allow them for employment to excel in a particular job.

“ PTs helped me to be more aware of the requirements of my future job because using the English language in real life tasks were really beneficial for me.” (Student 3)

“ By implementing the tasks I found an opportunity to use the ESP vocabulary items in real life related tasks so they reflected my learning.” (Student 5)

4.2.1.3 Developing awareness and social skills of the students. Public speaking is accepted one of the most significant forms of communication, so students need to master it in their educational life. Therefore, presentation skills of the students are required to be improved in schools because gaining that skill would be beneficial for them in their professional life in future. According to the views of the stakeholders, PTs could help the learners improve their public speaking and presentation skills which results in raising self-worth and self-confidence. On the basis of students' responses, presenting to the classmates helped them deal with the fear of public speaking which pushed the students to the next level. PTs helped them foster not only their presentation skills but also self-confidence of the students. That's why, some students were comfortable while talking in front of the other people. Some PTs boosted self-confidence of the students by making them face with their fears. Some of the students stated that they overcame the fear of speaking English thanks to PTs. On the contrary, some students started to have fun with using the target language at school environment.

Finding out the best way in order to enhance motivation of the students and engage them to the lesson is one of the significant issues raised by the teachers. It is widely recognized that students have affective filter when it comes to speaking. Students generally thought that the skill of speaking in the target language was a big challenge for them due to its interactive patterns. Based on the remarks obtained from the participants, it was clear that tasks were quite important in the sense that they contributed a lot to the self-confidence and speaking skills of the learners. They became aware of their potentials in language use. They learned how to manage their fear of speaking. What participants said about this issue in the focus discussion group is given below:

“ I have always found speaking English challenging, especially in front of other people. I felt very nervous every time that I needed to speak in front of others. I should admit that I felt nervous at the very beginning but seeing that I can speak with only minor faults in English, I felt very comfortable.”

“ Integrating different tasks to the assessment helped me broaden my perspective because I have been tested through traditional assessment up to now” (Student 7)

“ I felt more self-confident after performing some tasks in front of my classmates” (Student 6)

“ It was the first time I have done a presentation in English. I believe in myself more than before” (Student 9)

“ I think, speaking English during the PTs was very challenging.” (Student 5)

“ Doing presentations in front of the other people enhanced my self-confidence” (Student 3)

“ It allowed us to use the language meaningfully and creatively. Presentations encouraged to expand our speaking skills” (Student 8)

On the other hand, as the participants of the study were high school students performing some alternative tasks were quite different for them. As a requirement of the education system they are generally assessed through TA instead of PBA. That is to say, passing from memorizing the information to performing the knowledge they got during the classes is a big leap for them. Participants’ responses from the focus group discussion support the importance of PTs in increasing the sense of responsibility. Moreover, preparing and presenting the tasks helped them to increase the sense of responsibility. Participants’ responses from the focus group discussion support the importance of PTs in increasing the sense of responsibility. Additionally, PTs has contributed to them to be more autonomous learner because all tasks assigned were out of class tasks. Remarks of the students about taking responsibility of their own learning are given below:

“ I wanted to show my teachers how responsible I am by completing the PTs on time”(Student 3)

“ Actually, I do not like doing my homework but I spent more effort to complete PTs as I gave me more sense of responsibility. I gave importance to the tasks.” (Student 10)

“ When the subject couldn’t take my attention, I did not like spending time on the workload of that class so my mother helped me complete my homework. However, this was the first time I did all the tasks alone and it made me feel more successful.” (Student 5)

The results clearly showed that participants of the study generally had positive comments in the process of implementing PTs in vocational English class.

4.2.2 Challenges of using PBA tasks. During the implementation of performance based assessment, students experienced some disadvantages of being assessed through PTs.

1. Requiring longer time to complete
2. Putting more effort into PTs

4.2.2.1 Requiring longer time to complete. Analyzing the participants' remarks shows that some students had difficulty with spending enough time to the tasks as it took more time than studying for a traditional exam. Additionally, when other projects and exams were assigned, students indicated that they couldn't focus on some tasks. According to this theme, students do not favour PTs because they think that these tasks take too much time. What the students stated regarding time issues of the PTs are given below:

“ I do not like PTs because they take too long to complete.”(Student 8)

“ There should be alternative assessment methods in addition to traditional ones” (Student 2)

“ Preparing a well- designed performance task took so much time.” (Student 4)

“ We should implement both assessment methods in the next years .” (Student 3)

4.2.2.2 Putting more effort into PTs. Some students do not like PTs because they think that these tasks do not pay off for their time and effort. Some of the statements that suggest this are as follows:

“ I do not like PTs because I put a lot of effort on it but I never get good enough marks.” (Student 1)

“Studying for traditional assessment was easy because I just cover all topics the day before the exam but preparing a performance task required much effort in order to be more competent.” (Student 7)

All in all, we can conclude from the data gathered in vocational English class that being assessed via performance based assessment has some benefits and challenges. Students' motivation, awareness and social skills increased and language development fostered through the use of PTs. On the other hand, performing PTs required more time and effort when PTs were compared to the traditional assessment.

4.3 Findings of the Research Question 3. Do the students' perceptions of PBA tasks vary based on their proficiency level?

With the purpose of finding whether there was any statistically meaningful difference between the attitudes of the students towards PBA tasks according to students' proficiency level, Kruskal- Wallis Test was executed. The findings are presented in Table 12.

Table 12

Perceptions of the Students towards PBA according to Their Proficiency Level

		N	\bar{X}	Ss	H	Sd	P
Task Familiarity	Elementary	10	4,16	,33	6,770	2	,034*
	Pre-Intermediate	22	4,15	,77			
	Intermediate	20	4,41	,12			
Application	Elementary	10	3,66	,30	13,318	2	,001*
	Pre-Intermediate	22	3,83	,72			
	Intermediate	20	4,05	,13			
Scoring	Elementary	10	4,00	,43	7,551	2	,023*
	Pre-Intermediate	22	4,01	,79			
	Intermediate	20	4,30	,11			
Learning Outcomes	Elementary	10	3,74	,42	13,387	2	,001*
	Pre-Intermediate	22	4,12	,81			
	Intermediate	20	4,40	,22			
Program Consistency	Elementary	10	3,57	,30	2,737	2	,254
	Pre-Intermediate	22	3,61	,68			
	Intermediate	20	3,74	,12			
PBA Implementation	Elementary	10	3,82	,33	15,422	2	,000*
	Pre-Intermediate	22	3,96	,74			
	Intermediate	20	4,20	,10			

* $p < 0,05$

The Kruskal Wallis Test was applied to investigate whether there is a statistically significant difference in PBA implementation and stages of it the according to the proficiency levels of students who were participated in the study. As a result of the test, there is a statistically significant difference between the performance application

score and averages of the students' proficiency level because the 95% significance level is $p < 0,05$. It is seen that the consistency with the program phase of the performance application does not differ from the implementation stages as regards the level ($p > 0,05$). When averages are taken into the consideration, it is seen that the intermediate level students have the highest score average. On the other hand, the average values of the students in other levels are close to each other. In this case, it can be stated that the group that creates difference was students who were at the intermediate level.



Chapter 5

Discussion and Conclusion

This study aims to study how the ESP students perceive the use of performance-based assessment for the purpose of alternative assessment in vocational English class. In accordance with this purpose, both quantitative and qualitative data attributes have been used in this present study. The following parts of the study discuss the findings of each result regarding the research questions. Additively, this chapter presents conclusion of the study and also the recommendations for further research depending on the gathered data.

5.1 Discussion of Findings for Research Question 1. What are the students' perceptions of the PTs regarding task familiarity, application, scoring, learning outcomes, and program consistency?

In the light of data obtained from questionnaire, majority of the students are in favor of all phases of the implementation of the performance-based assessment. Performance-based assessment is a useful tool with the aim of practising English skills. Findings of the questionnaire concerning PTs illustrate that PTs are not capable enough to integrate all skills of the language. However, it is clear that contribution to the ESP vocabulary got the highest value which means that students agree on the improvement on their ESP lexical knowledge thanks to PTs. Additively, Yıldırım and Örsdemir (2013) state that PTs do not pay enough attention to some skills such as listening concerning the remarks of the students. In parallel with the findings of the questionnaire, it appears that the main issue in the observed PTs is related to listening and writing practice. Students perception on the contribution of PTs to the writing skill is not clear. In parallel with this result, Ozuslu (2018) puts forward that contrary to PTs' positive effect on productive skills, they slightly agreed with their contribution to the writing skills. As portfolio assessment, which was not in the scope of this study, was carried out in writing courses rather than PTs, this may direct students to disregard PTs' effect on their writing skills. However, language proficiency means not only the structures or vocabulary knowledge, but

also knowing how to use all language skills effectively (Yıldırım& Orsdemir, 2013). The results suggested here are consistent with those of Kemaloglu (2006) stating gaps in the development of speaking and listening and emphasizing importance of focusing on these skills as designing tasks and projects.

On the other hand, findings of the questionnaire show that students believe PTs help them to become more autonomus learners as they take charge of their own learning. Orsdemir (2018) states that the findings of the questionnaire reveal the fact that PTs help students enhance self- esteem and as they allocate responsibility. Moreoever, students enjoy preparing the tasks since they are eager to do them. These findings are in line with the results of study conducted by Zimbicki (2007) who points out the positive effect of alternative assessment methods on students' motivation.

Results of the questionnaire shed light that students mostly have positive attitude towards the use of PBA in vocational English class. Additionally, PTs should be used as supplementary assessment to TA. Ozuslu (2018) claims that the findings yet again may reveal necessity of longer time and experience for the learners so that they can analyze this innovative assessment method. Kızılkaya (2014) and Gökçen (2005) also focus on the need of students for more experience and time to show stronger attitude towards the alternative assessment practices. They may show tendency to use traditional assessment since they are not familiar with this new implementation. Cirit (2014) states that the students took advantage of the tasks but they still believe that the traditional assessment methods are necessary. Even students' perceptions are positive towards the tasks, they still have some hesitations on the implication of performance-based assessment as the only assessment tool.

5.2 Discussion of Findings for Research Question 2. What is the impact of PBA on students?

Based on the data collected by the participants, almost all students stated that performance- based assesement does not have only benefits but also drawbacks. The findings from the questionnaire and focus discussion group suggest that majority of the students are in favor of performance- based assessment even if there are some challenges of it. The findings revealed two major themes with a number of sub-themes that conceptually support them as follows. First theme is reaping benefits of PTs which

brought out sub-themes of increasing motivation, fostering language development, and developing awareness and social skills of the students. Second theme is experiencing challenges in using PTs. According to the data, requiring longer time to complete and requiring more effort are examples of challenges.

The data regarding participants attitudes towards the benefits of being assessed through PTs were collected by all the data collection tools, namely survey and focus group discussion. The obtained data from those tools enabled the researcher to find a response to the advantages of performance- based assessment. Initially, students' remarks show that main positive effect of PTs on students is motivation. Most of the students stated that PTs increased their motivation to complete the tasks. Chen (2007) supports this opinion by stating that in the present study, the students confirmed the account; the students realized higher tendency to perform the tasks despite time-consuming and demanding work. Additionally, making learning fun and interesting motivated students in the process of learning and preparing PTs. Motivation is believed to be one of the main reasons for learning a language in a successful way (Kesmen, 2018). Based on the student data, it is clear that a sense of positive motivation had developed in the students.

In terms of benefits of performance- based assessment, one of the mostly stated advantages of PTs is that they helped students improve their social skills and sense of responsibility. Kesmen (2018) sheds light to that fact by stating that they help learners take increasing responsibility for their own learning by encouraging them to be able to give reflection, self-evaluation, and action planning as a process for lifelong learning. Therefore, this study confirms the results when data gathered from the previous studies are taken into the consideration in that sense. In the line to these findings, Chen (2007) states that according to the students, the performance- based assessment is “stimulating,” “motivating,” “challenging,” and “promoting learning.”

Apart from the benefits of using PTs in vocational English class, there are some challenges faced according to the findings. When the findings obtained from the study are examined, the most common disadvantage stated by the students in all the data collection instruments is that it is time consuming and requires more effort when it compared to traditional assessment. Students spend almost two weeks in order to complete one PT; therefore, it seems they are used to easiness of traditional assessment which is one shot and quick. Orsdemir (2010) supports this result by stating that there are also students who believe that PTs do not pay off their time and effort.

Although the participants face some disadvantages of alternative assessment through PTs as pointed out in the focus discussion group interview and questionnaire, there are more advantages than the disadvantages proving that the participants tend to choose the performance- based assessment.

5.3 Discussion of Findings for Research Question 3. Do the students' perceptions of PBA tasks vary based on their proficiency level?

Findings of the study revealed that intermediate level students have the highest score average. In other words, students study at intermediate level have mostly positive attitude towards the PBA and its phases. PTs are defined as the activities that measure the higher level of thinking skills of the students. They are the alternative ways of assessment, which requires the students to create a product that demonstrates their knowledge and skills (Örşdemir, 2010). By showing knowledge and skills successfully students are likely to get high grades. It is clear that there is a consistency between students level of proficiency and academic achievement. Ross (2005) puts forwards that higher proficiency learners progressively get higher grade point averages in the EAP courses.

On the other hand, the average values of the students in other levels are close to each other. In this case, it can be stated that the group that creates difference was students who were at the intermediate level. These findings are in line with the results of study by Kırıkkaya& Vurkaya (2011) revealed that no significant difference was found between the pretest and post-test scores in the control groups in schools with 'medium' and ' low success' levels.

Islam and Irtisam (2018) states that as intake capacity of the students varies from level to level, the teachers should be aware of the fact that whether the students have the ability to ingest the ideas appropriately or not. Hence, use of alternative assessments can cause difficulties for students with lower proficiency. If students fail to follow the activities requirements, it can cause anxiety that distract the learning process. Islam and Irtisam (2018) argued that “the linguistic capacity is also a determiner here because it sometimes requires a minimum linguistic competence which beginners do not have” (p. 20).

It would not be too wrong to say that different tools of PBA can be applied for different level of the students such as portfolios for beginner, projects for intermediate,

and debates for advanced level. PBA tools focus on initiating approaches, making decisions, self-evaluation, collaborative learning, and authentic use of language which demand minimum linguistic ability which might not be present among the beginner level of students (Islam & Irtisam, 2018, p. 20).

5.4 Conclusions

This study was conducted to gain a deep insight into the perceptions of the students towards the use of PBA as an alternative assessment methods in a vocational English course. The data specifically focus on the students to get better understanding of their attitudes towards assessment types.

The quantitative data concerning students' attitude towards the PTs were collected through a five- point Likert- scale which had 5 phases; task familiarity, application, scoring, learning outcomes and program consistency and also focus discussion group used as qualitative data collection tool. It can be inferred that both data analysing tools show that the students positively perceive the alternative assessment methods, in the focus discussion group, most of the students showed a positive perception on using performance- based assessment methods and also by saying that the motivation is enhanced, social skills and self-awareness are improved as well as efficiency increased by the use of PTs although some students point out some challenges by stating that PTs take longer time to complete and need more effort than traditional assessment tools. Additively, students' perceptions towards PBAs vary according to their proficiency level.

Overall, when questionnaire and focus discussion group findings are taken into account, it can be concluded that students mostly have positive attitude towards the implementation of PTs in vocational English course. Also, it is stated that PTs should be combined with traditional assessment methods in vocational English course. It would not be too wrong to say that performance- based assessment as a form of alternative assessment contributed to testing students' knowledge in the long term in the ESP context.

To sum up, objective of the present study is to contribute to the going debate on alternative assessment tools by investigating the usage of these tools in the scope of assessment practices in vocational English classes.

5.5 Recommendations

Because of the reason that implementation of performance- based assessment in ESP classes is a new issue in Turkey, the studies in this field are not sufficient enough. Thus, carrying on further studies in this area would be beneficial for the sake of the future of the performance- based assessment. This study was a small scale as it occurred in one of the private schools in İzmir and also participants of this study were only students. For this reason, implementing this study in various vocational high schools and including the opinions of teachers and administrators may be more fruitful in terms of adding different dimensions to the other studies. Also, conducting the study with groups from different institutions which have ESP classes may provide a deeper understanding of the performance based assessment in ESP courses. In addition, the number of the students participating in the study was 52 but it could have been more. When further studies are done with more students, more reliable and generalizable data could be obtained. Lastly, students' previous experience with PBA tasks might be taken into the consideration in order to learn whether there is a difference on their perceptions of PBA or not.

REFERENCES

- Ahmed, M.K. (2014). The ESP Teacher: Issues, Tasks And Challenges. *English for Specific Purposes World*, 15, 1-33
- Aiguo, W. (2008). Reassessing the position of Aviation English: From a special language to English for Specific Purposes. *Ibérica: Revista de la Asociación Europea de Lenguas para Fines Específicos (AELFE)*, 15, 151-164.
- Ak, S. (2013). *The evaluation of performance tasks given at primary education according to students' teachers' and parents' opinions*. Master Dissertation, Çağ University, Mersin.
- Alderson, J. C. (2010). A survey of aviation English tests. *Language Testing*, 27(1), 51-72.
- Ardiyansah, T. (2019). Analysis of speaking assessment in ESP speaking class. *Journal Of English Teaching, Literature, and Applied Linguistics*, 3(1), 24-28.
- Arnold, D., Gramopadhye, A. K., Bingham, J., Master, R., & Quadros, N. (2000). Preparing the Aircraft Maintenance Technician for Tomorrow: Assessment of the New AMT Curriculum. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 44(22), 787.
- Ary, D., Jacobs, L. C., & Razavieh, A. (1996). *Introduction to research in education*. Fort Worth, Harcourt Brace College Publishers.
- Bachman, L. F. (1991). What Does Language Testing Have to Offer? University of California, Los Angeles. *TESOL Quarterly*, 25 (4), 671-704.
- Bailey, K. M. (1998). *Learning about language assessment: dilemmas, decisions, and directions*. Heinle & Heinle: US. Bailey, K. M. (1998). *Learning about language assessment: dilemmas, decisions, and directions*. Heinle & Heinle: US.
- Balliro, L. (1993). What kind of alternative? Examining alternative assessment. *TESOL Quarterly*, 27(3), 558-561.
- Barootchi, N., & Keshavarz, M. H. (2002). Assessment of achievement through portfolios and teacher-made tests, *Educational Research*, 44 (3), 279-288.
- Bas, G. (2011). Investigating the effect of project-based learning on students' academic achievement and attitudes towards English lesson. *The Online Journal of New Horizons in Education*, 1(4), 1-15.

- Baştürkmen, H. & Elder, C. (2006) The practice of LSP. In Alan DAVIES and Catherine EL - DER (eds.) *The handbook of applied linguistics* (pp. 672-694). Oxford: Blackwell
- Benavent, G.T., & Peñamaría, S.S. (2011). Use of authentic materials in the ESP classroom. *Universidad de Salamanca, Encuentro v20*, 89- 94.
- Berry, R., (2008), *Assessment for Learning*, Hong Kong: Hong Kong University Press.
- Brinke, D. J., Bruggen, J. V., Hermans, H., Burgers, J., Giesbers, B., Koper, R., & Latour, I. (2007). Modeling assessment for re-use of traditional and new types of Assessment. *Computers in Human Behaviour*, 23, 6, 2721-2741.
- Brown, H. D. (2004). *Language assessment: principles and classroom practices*. White Plains, NY: Pearson Education (pp.13).
- Brown, J. D., & Hudson, T. (1998). The alternatives in language assessment. *TESOL Quarterly*, 32 (4), 653-675.
- Brown, J.S. & Rodgers, T. (2004). *Doing second language research*. Oxford: Oxford University Press.
- Buluş Kırıkkaya, E. & Vurkaya, G. (2011). The effect of using alternative assessment activities on students' success and attitudes in science and technology course. *Educational Sciences: Theory & Practice*. 11. 885-1004.
- Büyüköztürk, S. (2002). *Sosyal bilimler için veri analizi el kitabı: İstatistik, araştırma deseni, SPSS uygulamaları ve yorum*. Ankara: Pegem A Yayıncılık.
- Can, E. (2017). *English teachers' classroom assessment practices and the ethics of classroom assessment practices*. Master dissertation, Çağ University, Mersin.
- Chen, Y. H. (2007). *Exploring the Assessment Aspect of Differentiated Instruction: College EFL Learners' Perspectives on Tiered Performance Tasks*. University of New Orleans Thesis and Dissertations. (UMI No.837).
- Cheng, L., Rogers, T., & Hu, H. (2004). ESL/EFL instructors' classroom assessment practices: Purposes, methods, and procedures. *Language Testing*, 21 (3), 360–389.
- Chun, M. (2010, March). "Taking teaching to (performance) task: Linking pedagogical and assessment practices." *Change: The Magazine of Higher Education*.
- Cirit, N. C. (2014). Assessing ELT Pre-Service Teachers via Web 2.0 Tools: Perceptions toward Traditional, Online and Alternative Assessment. *The Turkish Online Journal of Educational Technology*, 14(3), 9-19.
- Creswell, J. W. (2013). *Qualitative inquiry and research design: Choosing among five approaches* (3rd ed.). Thousand Oaks, CA: Sage.

- Creswell, J. W. (2014). *Educational Research: Task familiarity, Conducting and Evaluating Quantitative and Qualitative Research*, ed. vol. 4. United States of America: Pearson New International Edition.
- Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. Toronto: Holt, Rinehart, and Winston, Inc.
- Cumming, A. (2009). Language assessment in education: Tests, curricula, and teaching. *Annual Review of Applied Linguistics* 29, 90–100.
- Cutting, J. (2012). English for airport ground staff. *English for Specific Purposes*, 31, 3-13.
- Çakır, C., (2013). Standard Assessment and Alternative Assessment in English Language Teaching Program. *Gazi Eğitim Fakültesi Dergisi*, 33(3), 531-548.
- Çelik, S. (2015). *Investigating the effect of student response system supported think-pair-share pedagogy on preparatory school EFL students' vocabulary achievement*. Master dissertation, Middle East Technical University, Ankara.
- Demirdöken, G. (2019). *Ready for take-off: an aviation English needs analysis study in Turkey*. Master dissertation, Bahçeşehir University, İstanbul.
- Dietel, R. J., Herman, J. L., & Knuth, R. A. (1991). What does research say about assessment? NCREL, Oak Brook. Available online: http://www.ncrel.org/sdrs/areas/stw_esys/4assess.htm
- Dikli, S. (2003). Assessment at a Distance: Traditional vs. Alternative Assessments. *The Turkish Online Journal of Educational Technology*, 2(3), 13-19.
- Dörnyei, Z. (2003). *Questionnaires in second language research: Construction, administration and processing*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Eckert, D. (1997). *The use of simplified English to improve task comprehension for nonnative English speaking aviation maintenance technician students*. Retrieved from ProQuest (Access date: 05.10.2016).
- Enger, S.K. & Yager, R.E. (1998). *The Iowa Assessment Handbook*. (ERIC Document Reproduction Service No: ED424286).
- Fraenkel, J. R. & Wallen, N. E. (2006). *How to Design and Evaluate Research in Education*. McGraw-Hill International Edition, New York.
- Godfrey, J. R. (1998). Cunningham, George K. (1998). *Assessment in the classroom: Constructing and interpreting tests*. London: Falmer Press.
- Göçtü, R. (2013). *Comparison of traditional and alternative assessment in English Language Teaching in high schools (Georgian case)*. Ph.D dissertation, International Black Sea University, Tbilisi.

- Gökçen, R. A. (2005). *Instructors' and administrators' attitudes towards project work as an alternative assessment tool and as an instructional approach at Karadeniz Technical University School of Foreign Languages Department of Basic English*. Master dissertation, Bilkent University, Ankara.
- Gökhan, C. (2004). *Attitudes of teachers and students towards the assessment system at Hacettepe University basic English division and their opinions about alternative methods of assessment*. Master Dissertation, Bilkent University, Ankara.
- Hamayan, E. V. (1995). Approaches to Alternative Assessment. *Annual Review of Applied Linguistics*, 15, 212-226.
- Han, C. (2018). The Practical Study of ESP Theory in Professional Foreign Language Teaching. *2018 International Conference on Intelligent Transportation, Big Data & Smart City (ICITBS), Intelligent Transportation, Big Data & Smart City (ICITBS), 2018 International Conference on, ICITBS*, 386-389. <https://doi-org.libezproxy.iyte.edu.tr/10.1109/ICITBS.2018.00104>
- Hancock, C. R. (1994). Alternative assessment and second language study: What and why? *ERIC Digest*, ED 376 695, Retrieved May 28, 2014, from <http://eric.ed.gov/?id=ED376695>.
- Herman, J. L., Aschbacher, P. R., & Winters, L. (1992). *A practical guide to alternative assessment*. Alexandria, VA.: Association for Supervision and Curriculum Development.
- Huerta-Macias A. (1995). Alternative assessment: Responses to commonly asked questions. *TESOL Journal*, 5(1), 8-11.
- Hutchinson, T., & Waters, A. (1987). *English for specific purposes: A learning-centered approach*. Cambridge: Cambridge University Press.
- Hutchinson, T., & Waters, A. (1992). *English for Specific Purposes: A learning centered approach*. Cambridge: Cambridge University Press.
- Islam, N., & Ahmed, I. (2018). Thinking Unconventional: Alternative Assessment at Tertiary Level Education in Bangladesh. *International Journal of New Technology and Research (IJNTR)*, 4 (1), 18-23.
- Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher*, 33(7), 14–26.
- Kone, K. (2015). *The Impact of Performance-Based Assessment on University ESL Learners' Motivation*. Master dissertation, Minnesota State University, Mankato. All Theses, Dissertations, and Other Capstone Projects. (UMI No. 402).
- Kemaloglu, E. (2006). "Project Work: How Well Does It Work? Assessments of Students and Teachers about Main Course Project Work at Yıldız Technical

University School of Foreign Languages Basic English Department”, *Master Thesis*, Bilkent University, Ankara.

- Kesmen, A. (2018). *Portfolio unveiled: Attitudes of administrators, teachers, students and parents towards the use of language portfolios in a private primary school context*. Master Dissertation, Bahçeşehir University, İstanbul.
- Kızılkaya, B. (2014). *Students’ attitudes towards alternative assessment: A survey-based study on alternative assessment*. Master dissertation, İstanbul University, İstanbul.
- Law, B., & Eckes, M. (1995). *Assessment and ESL*. Peguis publishers: Manitoba, Canada.
- Libman, Z. (2010). Alternative Assessment in Higher Education: An experience in descriptive statistics. *Studies in Educational Evaluation*, 36 (1- 2), 62-68.
- Lincoln, Y.S., & Guba EG. (1985). *Naturalistic inquiry*. Beverly Hills, CA: Sage.
- Lomperis, A. E. (2010). *Issues in language policy for the labor force in developing countries*. In M. Gueldry (Ed.), *Consistent incorporation of professional terminologies into the world’s languages: The linguistic engine of a global culture* (pp. 311–341). Lewiston, NY: Edwin Mellen Press.
- Maykut, P., & Morehouse, R. (1994). *The Falmer Press teachers' library, 6. Beginning qualitative research: A philosophic and practical guide*. Falmer Press/Taylor & Francis, Inc.
- McNamara, T. (2001). Editorial: Rethinking alternative assessment. *Language Testing*, 18 (4), 329–332.
- Nasab, F. G. (2015). Alternative versus Traditional Assessment. *Journal of Applied Linguistics and Language Research*, 2(6), 165-178.
- Nation, I. S. P. (2013). *Learning Vocabulary in Another Language* (2nd Eds.). Cambridge: Cambridge University Press.
- O’Malley, J. M., & Pierce, L. V. (1996). *Authentic assessment for English language learners: practical approaches for teachers*. Massachusetts: Addison Wesley Publishing Company.
- Orsdemir, E. (2010). *Alternative assessment in Turkish primary EFL classrooms: an investigation into performance tasks*. Master dissertation, Çukurova University, Adana.
- Ozdamar, K. (2004), *Paket Programlar ile İstatistiksel Analiz*, Eskişehir: Kaan Kitabevi.

- Ozusu, N. (2018). *An evaluation of the performance assessment component of an EFL preparatory school program*. Master dissertation, Yeditepe University, İstanbul.
- Quansah, F. (2018). Traditional or Performance Assessment: What is the Right Way in Assessing Learners? *Research on humanities and social sciences*, 8, 21-24.
- Rahman, M. (2015). English for Specific Purposes (ESP): A Holistic Review. *Universal Journal of Educational Research*, 3(1), 24-31.
- Ross, S. (2005). The impact of assessment method on foreign language proficiency growth. *Applied Linguistics*, 26, 317-342.
- Saldana, J. (2016). *The coding manual for qualitative researchers*. London, UK: Sage.
- Saricoban, A., & E. Kuntas (2010). Assessing young learners. In H.H. Uysal & B. Haznedar (Eds.), *Handbook for teaching foreign languages to young learners in primary schools* (pp.351-374). Anı Yayıncılık.
- Seale, C. (1999). Quality in qualitative research. *Qualitative Inquiry*, 5(4), 465-478.
- Serpil, R. (2017). *The development of a technical word list and self- study material for aircraft maintenance students*. Master Dissertation, Anadolu University, Eskişehir.
- Simonson, M., Smaldino, S., Albright, M. and Zvacek, S. (2000). Assessment for distance education. *Teaching and Learning at a Distance: Foundations of Distance Education*. Upper Saddle River, NJ: Prentice-Ha.
- Shepard, L. (2008). The role of assessment in a learning culture. *Journal of Education*, 189(1/2), 95–106.
- Shenton, A. K. (2004). Strategies for Ensuring Trustworthiness in Qualitative Research Projects. *Education for Information*, 22, 63-75.
- Shohamy, E. (1995). Performance Assessment in Language Testing. *Annual Review of Applied Linguistics*, 15, 188-211. doi:10.1017/S0267190500002683
- Stiggings, R. (2001). *Student-involved classroom assessment*. (3rd ed.). Upper Sadle River, NJ: Merril-Prentice Hall.
- Stoica, A. M. (2006). Assessment methods for ESP. Retrieved from www.diacronia.ro/indexing/details/A5488/pdf.
- Streff, R. J. (2016). *A Qualitative Case Study of Strategies for Choosing and Evaluating Alternative Assessments in Online Higher Education*. Ph.D. dissertation, Walden University Walden, the U.S.

- Tan, S. W. (2016). *English for Occupational Purposes: Elastomer English* (Kent State University). Retrieved from https://etd.ohiolink.edu/pg_10?0::NO:10:P10_ETD_SUBID:115274
- Weigle, S. C. (2002). *Assessing writing*. Cambridge University Press.
- Wiggins, G. (1990). The case for authentic assessment. ERIC Digest. ERIC Document Reproduction Service No. ED 328 611. *Practical Research, Assessment & Evaluation*, 2(2).
- Winking, D. (1997). Critical issue: Ensuring equity with alternative assessments [online document]. NCREL (North Central Regional Educational Laboratory), Oak Brook: IL. Available online: <http://www.ncrel.org/sdrs/areas/methods/assment/as800.htm>
- Yıldırım, A. (2004). Student assessment in high school social studies courses in Turkey: Teachers' and students' perceptions. *International Review of Education*, 50(2), 157-175.
- Yildirim, R., & Orsdemir, E. (2013). Performance Tasks as Alternative Assessment for Young EFL Learners: Does Practice Match the Curriculum Proposal? *International Online Journal of Educational Sciences*, 5(3), 562-574.
- Zareva, A. (2019). Providing Feedback on the Lexical Use of ESP Students' Academic Presentations: Teacher Training Considerations. In *Research- publishing.net*. Research- publishing.net.
- Zhang, Z., & Burry-Stock, J. A. (2003). Classroom assessment practices and teachers' self-perceived assessment skills. *Applied Measurement in Education*, 16(4), 323- 342.
- Zimbicki, D. (2007), "*Examining the effects of alternative assessment on student motivation and self- efficacy*", Ph. D. dissertation, Walden University.

