

**OPINIONS OF ENGLISH TEACHERS IN STATE PRIMARY SCHOOLS ON  
THE TESTS THEY APPLY, THE EFFECT OF SBS ON THEIR TESTS AND THE  
PROBLEMS FACED**

**Pamukkale University  
Institute of Social Sciences  
Master of Arts Thesis  
English Language Teaching Department**

---

**Tuğba AKINCI**

**Supervisor: Asst. Prof. Dr. Selami OK**

**June, 2010**

**Denizli**

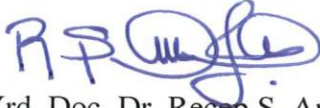
**YÜKSEK LİSANS TEZİ ONAY FORMU**

**İngiliz Dili Eğitimi** Anabilim Dalı, İngilizce Öğretmenliği Bilim Dalı öğrencisi **TUĞBA AKINCI** tarafından Yrd. Doç. Dr. Selami Ok yönetiminde hazırlanan “**Opinions of English Teachers in State Primary Schools on the Tests They Apply, the Effect of SBS on their Tests and the Problems Faced**” başlıklı tez aşağıdaki jüri üyeleri tarafından **24/06/2010** tarihinde yapılan tez savunma sınavında başarılı bulunmuş ve Yüksek Lisans Tezi olarak kabul edilmiştir.



Yrd. Doç. Dr. Selami Ok

Jüri Başkanı (Danışman)



Yrd. Doç. Dr. Recep Ş. Arslan

Jüri Üyesi



Yrd. Doç. Dr. Mehmet Ali Çelikel

Jüri Üyesi

Pamukkale Üniversitesi Sosyal Bilimler Enstitüsü Yönetim Kurulu'nun **15/07/2010** tarih ve ...**11/01**... sayılı kararıyla onaylanmıştır.

**Doç. Dr. Bilal SÖĞÜT**  
Müdür

Bu tezin tasarımı, hazırlanması, yürütülmesi, arařtırmalarının yapılması ve bulgularının analizlerinde bilimsel etiĐe ve akademik kurallara özenle riayet edildiĐini; bu alıřmanın doĐrudan birincil ürünü olmayan bulguların, verilerin ve materyallerin bilimsel etiĐe uygun olarak kaynak gösterildiĐini ve alıntı yapılan alıřmalara atfedildiĐini beyan ederim.

İmza :  
ÖĐrenci Adı Soyadı : TuĐba AKINCI

## ACKNOWLEDGEMENTS

The completion of this study would not have been possible without the help of many people. First, my sincere appreciation goes to my thesis supervisor, Asst. Prof. Dr. Selami OK. I would not have been able to accomplish my study without his support, guidance and encouragement.

My thanks also go to my teachers from the department, Asst. Prof. Dr. Turan PAKER and Asst. Prof. Dr. Recep Şahin ARSLAN for their insightful feedback and support during this learning process. I would also like to express my gratitude to Asst. Prof. Dr. Mehmet Ali ÇELİKEL.

I'm indebted to the teachers who participated in the study for helping me with the data collection in Kartal, İstanbul. This study would not have been possible without their contributions.

I'm deeply grateful to my family for their love and belief in me. My parents and brothers have always been there for me when I needed them. Without their continuous support and encouragement, I could not have achieved this goal.

Many thanks go to my beloved husband, Mehmet AKINCI, who challenged, advised and entertained me throughout this entire process. A thank-you also goes to my husband's family for their kindness and understanding.

**ÖZET**  
**DEVLET İLKÖĞRETİM OKULLARINDA ÇALIŞAN İNGİLİZCE**  
**ÖĞRETMENLERİNİN UYGULADIKLARI SINAVLAR HAKKINDAKİ**  
**GÖRÜŞLERİ, SBS' NİN BU SINAVLAR ÜZERİNDEKİ ETKİLERİ VE**  
**KARŞILAŞILAN SORUNLAR**

Akıncı, Tuğba  
 Yüksek Lisans Tezi, İngiliz Dili Eğitimi Anabilim Dalı  
 Tez Danışmanı: Yard. Doç. Dr. Selami OK  
 Haziran 2010, 112 Sayfa

**Bu çalışma devlet ilköğretim okullarında çalışan İngilizce öğretmenlerinin ölçme hakkındaki görüşlerini ve karşılaşılan sorunları incelemek amacıyla yapılmıştır. Çalışma aynı zamanda Seviye Belirleme Sınavı' nın öğretmenlerin hazırladıkları sınavlar üzerindeki etkilerini de araştırmaktadır.**

**Bahsi geçen amaçlar göz önüne alınarak bir araştırma düzeni hazırlanmıştır. Pilot çalışmayı takiben, esas çalışma gerçekleştirilmiştir. Veriler araştırmacı tarafından hazırlanan bir anket ve görüşme tekniği aracılığıyla toplanmıştır. Bu çalışmaya İstanbul, Kartal ilçesindeki 80 İngilizce öğretmeni katılmıştır. Çalışma, 2009-2010 eğitim öğretim yılında yapılmıştır. Anket çalışmasından elde edilen veriler SPSS (12.00) Sosyal Bilimlerde İstatistiksel Veri Analizi programı ve Microsoft Office 2003 Excel programlarıyla değerlendirilmiştir. Anketin bazı kısımlarından ve yapılan görüşmelerden elde edilen veriler ise nitel analiz gerektirmiştir.**

**Her iki veri toplama aracından elde edilen sonuçlar, öğretmenlerin konuşma ve dinleme becerilerinin öğretime verdikleri önem derecesi ile bu iki becerinin ölçülmesine verdikleri önem derecesi arasında çelişkiler olduğunu göstermiştir. Özellikle konuşma ve dinleme becerilerinin ölçülmesine gereken önemin verilmemesini etkileyen faktörler ise kalabalık sınıflar, araç-gereç ve zaman yetersizliği, öğrencilerin yetersiz seviyeleri olarak belirlenmiştir. Beceri ve alt beceriler arasında öğretime ve ölçülmesine en fazla önem verilen kelime bilgisi olmuştur. Kelime bilgisinin ölçülmesinde en etkili faktörler ise kelime bilgisinin derste öğretilmesi, SBS'nin kelime bilgisini ölçen sorular içermesi ve müfredatın kelime bilgisi içermesi olmuştur. Bunun yanı sıra, sonuçlar öğretmenlerin sınavlarında en fazla kullandığı soru tipleri boşluk doldurma, eşleştirme ve çoktan seçmeli sorular olduğunu ve öğretmenlerin büyük çoğunluğunun sınavlarında görsel malzeme kullanmayı tercih ettiklerini göstermiştir. Ayrıca sonuçlar, Milli Eğitim Bakanlığı tarafından hazırlanan SBS'nin öğretmenlerin öğretim ve ölçme uygulamaları üzerinde etkileri olduğunu göstermiştir. Okulda kullanılan ders kitaplarının öğrencileri SBS'ye hazırlamak için yetersiz olduğu ve bu yüzden öğretmenlerin büyük çoğunluğunun ek kaynak kullandığı sonuçlarına da varılmıştır.**

**Anahtar Kelimeler:** Sınavlar, Ölçme, Dil Becerileri ve Alt Becerileri, SBS

**ABSTRACT**  
**OPINIONS OF ENGLISH TEACHERS IN STATE PRIMARY SCHOOL ON**  
**THE TESTS THEY APPLY, THE EFFECT OF SBS ON THEIR TESTS AND THE**  
**PROBLEMS FACED**

Akıncı, Tuğba  
M.A Thesis in ELT  
Supervisor: Asst. Prof. Dr. Selami OK  
June 2010, 112 Pages

The present study was conducted to examine the opinions of English teachers working in state primary schools on their practices of testing and the problems they encounter. It also aimed to explore the effects of SBS on English teachers' test.

Considering the mentioned aims, the research was conducted through survey methodology. Following the pilot study, the main study was carried out. Data collected through a questionnaire and an interview which were prepared by the researcher. 80 English teachers in Kartal, İstanbul participated in this study. The study was conducted in 2009- 2010 academic year. The data obtained from the questionnaire were analyzed with SPSS 12.00 frequency analysis and Microsoft Office 2003 Excel programs. The data gathered from some parts of questionnaire and the interview needed qualitative analysis.

The results of both data collection instruments indicated that there were contradictions between the importance given to teaching speaking and listening and the importance given to testing these two skills. Factors affecting testing speaking and listening skills, which were given the least importance in testing, were determined as crowded classes, lack of equipment and time, students' low level of proficiency in these skills. Among the language skills and subskills, the most importance was given to teaching and testing vocabulary. The most effective factors in teaching and testing vocabulary were the fact that teachers teach vocabulary in class, SBS includes vocabulary questions and the curriculum covers vocabulary. In addition, results indicated that teachers mostly use gap-filling, matching and multiple choice items and great majority of the teachers make use of visuals in their tests. The results also showed that SBS, prepared by Ministry of Education had effects on teachers' teaching and testing practices. It was also concluded that the textbooks used at school are insufficient to prepare the students for SBS; hence, most of the teachers use supplementary materials.

**Keywords:** Tests, Testing, Language skills and subskills, SBS.

## TABLE OF CONTENTS

ACKNOWLEDGEMENT .....	ii
ÖZET .....	iii
ABSTRACT .....	iv
TABLE OF CONTENTS .....	v
LIST OF TABLES .....	viii
LIST OF FIGURES .....	ix
LIST OF ABBREVIATIONS .....	x

### CHAPTER ONE INTRODUCTION

1.1. Background of the Study.....	1
1.2. Statement of the Problem .....	3
1.3. The Aim and Significance of the Study .....	4
1.4. The Research Questions .....	4
1.5. Assumptions and Limitations of the Study .....	5
1.6. Outline of the Study .....	5
1.7. Operational Definitions .....	6

### CHAPTER TWO LITERATURE REVIEW

2.1. Definitions of Assessment and Testing .....	8
2.2. Historical Background of Language Testing .....	9
2.3. Purposes of Language Testing .....	11
2.4. Principles of Testing .....	12
2.4.1. Practacity .....	12
2.4.2. Validity .....	12
2.4.3. Reliability .....	14
2.4.4. Authenticity .....	16
2.4.5. Washback .....	16
2.5. Classification of Language Tests .....	17
2.5.1. Classification of Language Tests According to Their Content .....	17
2.5.1.1. Proficiency Tests .....	17
2.5.1.2. Achievement Tests .....	18
2.5.1.3. Language Aptitude Tests .....	18
2.5.1.4. Direct versus Indirect Tests .....	19
2.5.1.5. Discrete-point versus Integrative Tests .....	19
2.5.2. Classification of Language Tests According to Their Frame of Reference.....	20
2.5.2.1. Norm-referenced Tests .....	20
2.5.2.2. Criterion-referenced Tests .....	20
2.5.3. Classification of Language Tests According to Their Scoring Procedure .....	20
2.5.3.1. Objective Tests versus Subjective Tests .....	21

2.6. Standardized Testing .....	21
2.7. Testing Language Skills .....	22
2.7.1. Testing Listening Comprehension.....	22
2.7.2. Testing Speaking .....	24
2.7.3. Testing Reading comprehension .....	26
Comprehension.....	
2.7.4. Testing Writing .....	27
2.7.5. Testing Vocabulary .....	29
2.7.6. Testing Grammar .....	30
2.8. Types of Test Items .....	31
2.8.1. Multiple Choice Items .....	31
2.8.2. Binary Choice Items .....	32
2.8.3. Gap-filling Items .....	32
2.8.4. Matching Items .....	33
2.8.5. Cloze Items .....	33
2.8.6. Paragraph Writing .....	34
2.9. The Use of Visuals in Language Tests .....	34
2.10. English Language Teaching Curriculum in Comparison with SBS .....	35

### **CHAPTER THREE METHODOLOGY**

3.1. Research Design .....	37
3.2. Participants .....	40
3.3. Data Collection Instruments .....	41
3.3.1. Questionnaire .....	41
3.3.1.1. Pilot Study .....	42
3.3.1.1.1. Piloting Procedure .....	42
3.3.1.1.2. Results of the Pilot Study .....	43
3.3.1.2. Content of the Questionnaire .....	44
3.3.2. Interview .....	46
3.4. Data Analysis .....	48

### **CHAPTER FOUR RESULTS AND DISCUSSION OF THE FINDINGS**

4.1. Introduction .....	49
4.2. Results .....	49
4.2.1. Findings on the First Resarch Question .....	50
4.2.1.1. Importance Given to the Language Skills /Subskills .....	50
4.2.1.2. Language Skills/Subskills Tested .....	53
4.2.2. Findings on the Second Research Question .....	54
4.2.2.1. Vocabulary .....	54
4.2.2.2. Grammar .....	55
4.2.2.3. Reading comprehension Comprehension.....	57
4.2.2.4. Writing .....	59



4.2.2.5.	Listening Comprehension.....	61
4.2.2.6.	Speaking .....	63
4.2.3.	Findings on the Third Research Question .....	65
4.2.3.1.	The Importance Given to the Item Types by the Participants .	65
4.2.3.2.	The Question Types Preferred to Test Skills/Subskills by the Participants .....	67
4.2.4.	Findings on the Fourth Research Question .....	71
4.2.5.	Findings on the Fifth Research Question .....	73
4.2.5.1.	The Consistency between the Textbooks and SBS .....	73
4.2.5.2.	Supplementary Materials Used for SBS .....	75
4.2.6.	Findings on the Sixth Research Question .....	76

## **CHAPTER FIVE CONCLUSION**

5.1.	Introduction .....	79
5.2.	Overview of the Study .....	79
5.3.	Conclusion .....	80
5.4.	Implications of the Study .....	83
5.5.	Suggestions for Further Research .....	85
	REFERENCES .....	86
	APPENDICES .....	93
	Appendix 1 Questionnaire .....	93
	Appendix 2 Interview Questions .....	100
	Appendix 3 Reliability Outputs of Pilot Study .....	101
	Appendix 4 Reliability Outputs of Main Study .....	102
	Appendix 5 2009 SBS English Questions for 6 <sup>th</sup> , 7 <sup>th</sup> and 8 <sup>th</sup> Grades ....	103
	CURRICULUM VITAE .....	113

## LIST OF TABLES

Table 3.1.	Gender of the Participants .....	40
Table 3.2.	University Departments Participants Graduated from .....	40
Table 3.3.	Means of the Participants' Work Experience .....	41
Table 3.4.	Gender of the Interviewees .....	47
Table 3.5.	University Departments Interviewees Graduated from .....	47
Table 3.6.	Means of the Interviewees' Work Experience .....	48
Table 4.1.	Percentages of Skills/Subskills Participants Give Importance in Their Teaching .....	50
Table 4.2.	Means of Skills/Subskills Participants Give Importance in Their Teaching .....	51
Table 4.3.	Percentage of the Participants Testing or not Testing Language Skills .....	53
Table 4.4.	Factors and Percentages of Testing Vocabulary .....	54
Table 4.5.	Factors and Percentages of Testing Grammar .....	56
Table 4.6.	Factors and Percentages of Testing Reading Comprehension .....	57
Table 4.7.	Factors of not Testing Reading Comprehension.....	58
Table 4.8.	Factors and Percentages of Testing Writing .....	59
Table 4.9.	Factors and Percentages of not Testing Writing .....	59
Table 4.10.	Factors and Percentages of Testing Listening Comprehension .....	61
Table 4.11.	Factors and Percentages of not Testing Listening Comprehension ...	61
Table 4.12.	Factors and Percentages of Testing Speaking .....	63
Table 4.13.	Factors and Percentages of not Testing Speaking .....	64
Table 4.14.	Percentages of Question Types Participants Give Importance in Their Exams .....	66
Table 4.15.	Mean Scores of Question Types Preferred .....	66
Table 4.16.	Question Types Preferred to Test Vocabulary and Their Percentages	68
Table 4.17.	Question Types Preferred to Test Grammar and Their Percentages ...	68
Table 4.18.	Question Types Preferred to Test Reading comprehension and Their Percentages ....	69
Table 4.19.	Question Types Preferred to Test Writing and Their Percentages .....	69
Table 4.20.	Question Types Preferred to Test Listening comprehension and Their Percentages .....	70
Table 4.21.	Question Types Preferred to Test Speaking and Their Percentages ...	70
Table 4.22.	Factors and Percentages of Using Visuals in Exams .....	71
Table 4.23.	Factors and Percentages of not Using Visuals in Exams .....	72
Table 4.24.	EFL Teachers' Opinions on the Sufficiency of the Textbooks for SBS.....	73
Table 4.25.	Percentages of Participants' Supplementary Materials Usage .....	75

**LIST OF FIGURES**

Figure 2.1. Tests and Assessment .....	8
--	---

**LIST OF ABBREVIATIONS**

<b>EFL</b>	English as a Foreign Language
<b>ELT</b>	English Language Teaching
<b>MOE</b>	Ministry of Education
<b>SBS</b>	Seviye Belirleme Sınavı (Level Identification Exam)

## **CHAPTER ONE**

### **INTRODUCTION**

#### **1.1. Background of the Study**

There has been a growth in the attention paid to testing to improve the quality of education (Herman et al. 1990). It is one of the essential parts of language teaching process. In this process teachers can be said to have a crucially important role since they choose and shape the way to go. Hamp-Lyons (2000: 580) lays a strong emphasis on the role of teachers in testing and suggests that “the vast majority of people, who design, prepare, administer and mark language tests are teachers”. While doing the heavy work of testing, teachers’ opinions affect their choice; however, most of them are not fully aware of how effective their preference is in teaching language. With respect to opinions, Wright (1987) claims that they have profound influence on the whole educational process. In addition, Karavas (1996) states that “teachers’ educational attitudes and theories, although in many cases unconsciously held, have an effect on their classroom behaviors, influence what students actually learn, and are a potent determinant of teachers’ teaching style” (p.188). Furthermore, Williams and Burden (1997) argue that teachers’ opinions are far more influential than their knowledge in their actions.

In spite of the fact that the effects of language teachers’ opinions on testing have not been investigated, some research have been conducted on statewide testing and opinions of teachers’ on this testing process ( Brown, 1992; Jet & Schafer, 1993; Cimbricz & College, 2002; Abrams et al.,2003). With the aim of exploring the opinions of teachers on statewide tests, Brown (1992) conducted a study and his findings indicated that teachers preferred employing traditional methods rather than applying whole language, cooperative learning and higher order thinking activities in order to be successful in statewide test. In addition to

his findings, Brown diverts the attention to the standardized tests and stresses that it is a debatable issue whether standardized tests affect curriculum and classroom instructions or not.

Dorr-Bremme & Herman (1986, in Smith et al., 1986) conducted a study on both internal (classroom tests) and external (state-wide tests) testing. The results of the study indicated that internal tests helped teachers to support instruction and evaluation of the learners, on the other hand, external tests regardless of being norm-referenced or criterion-referenced did not have the same implications. Darling-Hammond & Wise (1985, in Brown, 1992: 7) conducted a survey on the effects of standardized tests on teachers. The results indicated that standardized tests shaped more than half of the teachers' opinions in the class. Since those teachers changed the 'curricular emphasis' and taught learners how to be successful in the test, they could not allocate sufficient time to other materials. In addition to Darling-Hammond & Wise, Abrams et al. (2003) emphasize that teachers generally focus on what is tested. In an attempt to help learners achieve higher scores, teachers feel under constraint specifically in high-stakes tests, which can reduce the quality of education. In their survey, Abrams et al. (2003) revealed that teachers designed their classroom assessment in parallel with the high-stake state tests.

In addition to the powerful effects of teachers' opinions on testing, there are other factors playing significant roles in testing. One of these factors is considered to be the problems encountered in the process of testing. McNamara (2000) mentions several constraints in testing such as financial situation, lack of technology for listening comprehension and speaking, test security whether test content will be secure or not till the test date. Besides these problems, Davies (1990) states that time limitation is a restriction to test desired behavior since testing time duration is not enough to test the whole material taught and physical and psychological condition the tester in is another issue in testing.

## 1.2. Statement of the Problem

Language is a whole with all of its components such as listening, speaking, reading comprehension, writing, grammar, vocabulary and so on. All of the skills and subskills should be given emphasis in teaching and testing. In the second part of the Regulation on Teaching Foreign Languages in Turkey, prepared by Ministry of Education, the fifth item defines the aim of teaching foreign language:

“In formal, informal and distance education institutions, the aim of foreign language education is, in accordance with the general aims and fundamentals considering aims and levels of schools and institutions, in the foreign language taught, the individuals are enabled to gain a) listening comprehension, b) reading comprehension, c) speaking, d) writing skills, to communicate in the language s/he learnt and to develop positive attitudes towards foreign language education.” (Mevzuat Bankası, Milli Eğitim Bakanlığı Yabancı Dil Eğitimi ve Öğretimi Yönetmeliği, İkinci Kısım, Madde 5)

However, teachers do not seem to cover all the components whether in class or in the exam or both. Generally language components such as grammar, vocabulary and reading comprehension overweigh the other skills. Since to teach and test listening comprehension, speaking, writing need tremendous work, in most cases teachers’ opinions determine the process of teaching and testing.

Even though a number of studies have been conducted on the opinions of students toward testing so far, the teachers’ opinions on this major element of language teaching have not been paid attention a lot. Since many language teachers prepare, administer and evaluate their test on their own in Turkey, studies examining opinions of these teachers on testing gain increasing importance in literature.

Besides teachers opinions, SBS (Level Identification Exam) is another factor which affects the tests which English teachers prepare. SBS is a high stake, standardized exam, which has been administered at end of each academic year, to primary school students at 6<sup>th</sup>, 7<sup>th</sup>, and 8<sup>th</sup> grade students. In the test, each grade is given different number of English language questions. 6<sup>th</sup>, 7<sup>th</sup> and 8<sup>th</sup> grade students are asked to 13, 15, 17 questions

respectively in English. Therefore; attitudes of the students have changed toward English and teachers tend to teach and test in parallel with SBS format.

### **1.3. The Aim and Significance of the Study**

The aim of this study is to highlight the opinions of EFL teachers in state primary schools on their practices of testing and the effect of SBS on these tests and the problems faced. More specifically, the study aims at investigating the language skills mostly taught and tested at schools, the problems encountered during the process of testing, types of test items preferred to use and the use of visuals in the exams. In addition, the study aims at revealing the opinions of EFL teachers on the influence of SBS on their examinations, the sufficiency of the textbook for preparation to SBS, the use of supplementary materials for SBS.

Firstly, since there is little or no study conducted about the opinions of teachers on testing, the results will contribute to the field of testing at primary state schools by addressing the missing knowledge in the current literature regarding EFL teachers' classroom test practices, the problems they face and the effect of SBS on these tests. Secondly, this study was conducted in Turkish context and it will provide useful information for educators and administrators in Turkey for revising and developing EFL teacher training programs dealing with testing.

### **1.4. The Research Questions**

In order to achieve the aims of the study, we have tried to answer the following research questions:

1- Which language skills/subskills do state primary school EFL teachers test in their exams at school?

2- Which factors influence EFL teachers' opinions and their practices of testing?



- 3- Which question types do EFL teachers prefer to use in their exams at school?
- 4- To what extent do EFL teachers make use of visuals in their exams?
- 5- What are the opinions of state primary school EFL teachers on SBS?
- 6- What are the influences of SBS on the tests EFL teachers' apply?

### **1.5. Assumptions and Limitations of the Study**

This study was conducted at a local level and it is assumed that the participants represent the target population.

The questionnaire was distributed to teachers and the interview were held in Turkish in order to avoid confusion and to help the participants to understand the questionnaire items and the interview questions.

The sample population of this study is limited to 80 English teachers for the questionnaire and 8 English teachers for the interview, who work in state primary schools in Kartal, Istanbul.

### **1.6. Outline of the Study**

This study includes five chapters. Chapter One introduces the subject of the thesis, background of the study, statement of the problem, aim and significance of the problem, the research questions, assumptions and limitations of the study and operational definitions.

Chapter Two consists of the related literature on testing. After definitions of testing and assessment, the chapter continues with a historical background of language testing. Next, the literature on purposes of language testing, principles of language testing, classification of language tests, standardized tests, testing language skills/subskills and types of test items, the use of visuals in language tests are examined. The chapter ends with the comparison of the English Language Teaching Curriculum in Turkey and the English component of SBS.

Chapter Three introduces the methodology of the study such as research design, participants, data collection instruments and data analysis.

Chapter Four analyses the results of the data collection instruments; the questionnaire and the interview.

Chapter Five presents an overview of the study, conclusion and discussion of the findings, some implications and suggestions for further study.

### **1.7. Operational Definitions**

**Exam:** Exam is an exercise designed to examine progress or test qualification or knowledge. In this study ‘exam’ and ‘test’ are used interchangeably.

**Testing:** Testing is an “ administrative procedure that occurs at identifiable times in a curriculum when learners muster all their faculties to offer peak performance” (Brown, 2003:4).

**Language Subskills:** Language elements are grammar, vocabulary and pronunciation.

**Language Skills:** Language skills are categorized into four: listening comprehension, speaking, reading comprehension and writing.

**SBS (Level Identification Exam):** Ministry of Education gives an exam at the end of each academic year for the students at 6<sup>th</sup>, 7<sup>th</sup> and 8<sup>th</sup> grades, to test whether they successfully gained what is aimed in the curriculum that year. The exam questions are based on the gains and prepared to test interpretation, analysis, critical thinking, estimating the results and problem solving in obligatory classes such as Turkish, Mathematics, Science, Social Sciences and English (Tebliğler Dergisi, Kasım 2007). It is not an obligatory exam,

but it is suggested that the exam is important for placement in secondary schools which accept students who take this exam.

## CHAPTER TWO

### REVIEW OF LITERATURE

For several decades, many studies have been conducted in the field of English language teaching (ELT) in order to offer explanations regarding different aspects of language teaching and its assessment. In this chapter, the literature focusing on the definitions of assessment and testing, a historical background of language testing, the purposes of language testing, principles of language testing, classification of language tests, standardized tests, testing language skills/subskills and types of test items, the use of visuals in language tests, English Language Teaching Curriculum in comparison with SBS will be reviewed.

#### 2.1. Definitions of Assessment and Testing

Though assessment and testing are thought to be the same words and are used interchangeably, in fact they are not. According to Brown (2003:4), tests are administered at certain times in a curriculum and learners know that their responses are being measured and evaluated. As for assessment, it is a continuous process that includes a much wider domain. In contrast to testing, assessment can occur at any time when students answer a question, comment on a specific topic or make an effort to produce a word, phrase or structure. In addition, assessment can be made subconsciously by teachers. Brown (2003:5) shows the relationship between assessment and testing as in the Figure 2.1.

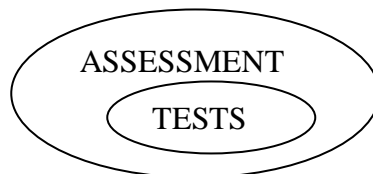


Figure 2.1. Tests and Assessment

As seen in the Figure 2.1., testing can be considered as subset of assessment. Using test is one of the procedures that the teachers can follow in order to assess the students' overall performance.

Popham (2003) differentiates the tests and the assessment by stating that the former is traditional (e.g. paper and pencil forms) whereas the latter is both traditional and communicative (e.g. portfolio products). He also suggests that if both traditional and communicative methods are combined in a test, the term 'test' can replace the term 'assessment' or vice versa. Coombe et al. (2007) summarize the difference between testing and assessment and describe assessment as “ all types of measures to evaluate students progress” while “ tests are a subcategory of assessment” (p.xv).

## **2.2. Historical Background of Language Testing**

Throughout history of language teaching and testing, the way of teaching shaped to the way to test the language. There has been several approaches to language teaching so has been to language testing. In attempt to group these approaches in language testing, Spolsky (1978) categorizes language testing into three: pre-scientific period, psychometric-structuralist period and integrative-sociolinguistic period. During the pre-scientific period, there were no kinds of 'statistical matters' such as validity, reliability, objectivity or other traits of language testing and also there were no rubrics or set criteria. Instead, there were experienced teachers taking the responsibility for not only teaching but also administering the tests and interpreting their results. Therefore, they might have suffered from low reliability. At this period, Clark (1983) emphasizes the use of grammar translation methods in both teaching and testing. Methods such as translating passages from target language to mother tongue or vice versa, grammar and culture of the target language were popular.

In psychometric structuralist period, in contrast to pre-scientific period ideas regarding that testing should be precise, objective, reliable, valid and scientific emerged and

made vital contributions to the development of testing (Spolsky,1978; Shohamy & Reves,1985). Spolsky (1978) focuses on the benefits of Robert Lado's studies for the development of this period. Spolsky(1978:8) argues "...for he accepts the testers right to establish kinds of tests and methods of judging validity and reliability even while insisting on the responsibility of the linguist to decide what is to be tested". Besides, the standardized tests were developed, which was the most remarkable result of Lado's studies. In this period, the tests included some elements of the language such as sounds, words and structures without a context; that is, only a specific part of the language was tested. Oral tests took place and they consisted of only the repetition of words and sentences or pattern questions to pattern answers. The tests were conducted in the language laboratories with machines, which was far from real life, to record the words, sentences or the answers (Shohamy & Reves, 1985). The tests used in this period were called discrete-point tests (Stansfield, 2008).

Discrete-point approach fails to cover the overall language ability since the test only measures limited knowledge and requires a de-contextualization leading confusion for test takers. Discrete-point tests did not serve for communicative purposes and couldn't have revealed the communicative ability of the learners (Brindley, 2001). The constraints mentioned led to a new period which laid emphasis on communication, context and authenticity is named as integrative-sociolinguistic stage by Spolsky (1978). While psycholinguists concern with integrative part of the language by stating that the language cannot be separated into discrete parts, rather it is a whole, sociolinguists propose the idea of "communicative competence" (p.9). Brindley (2001) claims that this trend increased integrative tests such as cloze and dictation which learners needed to reconstruct the meaning of spoken or written texts through use of linguistic and contextual knowledge. Weir (1990) claims that these integrative tests were indirect in nature and they were not test learners' performance ability directly. One of the leading scholars of integrative era Oller had a hypothesis which is known as "*unitary competence hypothesis*". It was based on his findings which reflect the view that "performance on a whole range of tests depends on the same underlying capacity in the learner - the ability to integrate grammatical, lexical, contextual, and pragmatic knowledge in test performance" (McNamara 2000:15). Although integrative tests required the gain of controlling several language skills at the same time,

they were indirect. The situation led defenders of communicative language testing to discuss that even though indirect tests had reliability and concurrent validity, other types of validity were under suspicion (Weir, 1990). Scholars were gaining more insight into language testing, the need for communicative testing were raising and according to Brown (2003) by the mid-1980s, the language-testing field had begun to focus on designing communicative language testing tasks. According to Canale and Swain (1980:4), communicative competence includes linguistic competence (knowledge of linguistic forms), sociolinguistic competence (the ability to use language appropriately in contexts), discourse competence (coherence and cohesion), and strategic competence (knowledge of verbal and non-verbal communicative strategies). In addition, Canale and Swain(1980) turn attention to the principles of communicative approach. They argue that the elements of communicative competence should not overweigh each other; needs of learners should be taken into consideration, the learner should have the chance to interact, the learning stages and steps in teaching should be well-planned. They emphasize that the communicative tests should seek for not only knowledge and competence but also the ability to perform these in a context. Furthermore, Bachman (1990, in Brown, 2003:10) comes up with a model of language competence being composed of “organizational and pragmatic competence, respectively subdivided into grammatical and textual component, and into illocutionary and sociolinguistic components”.

### **2.3. Purposes of Language Testing**

Testing in general has a variety of purposes. Henning (1987) examines language testing purposes from the teachers’ point of view and states that the purposes are to diagnose and give feedback; screen and select learners; place them; evaluate a program or provide research criteria. Bachman and Palmer (1996) lay emphasis on two purposes of language testing, the first one of which is to make inferences about language ability; and the other is to make decisions based on those inferences.

## **2.4. Principles of Language Testing**

In this subsection, principles of language testing such as practicality, validity, reliability, authenticity, and washback are introduced.

### **2.4.1. Practicality**

In preparing, conducting and scoring, practicality is one of the most essential principles in testing. Practicality is about the content, objective, administration, scoring of a test. The environment which the test will be conducted, the readiness of the equipment which will be used in the exam, enough copy for the testees and the cost of the test are the issues of practicality (Valette, 1987). McMillan (2007) emphasizes that practicality is the combination of many factors. Firstly, teacher familiarity with the testing method is important for practicality. Teachers should have enough knowledge about the test method, the appropriateness of the method to the learning objectives, the pros and cons of the technique, the administration process, scoring and interpretation. Another factor is enough time for the test preparation, administration and scoring. Time should be well-planned according to the test-method, test items and test takers. Thirdly, easy scoring and interpretation is a significant factor. Scoring and interpretation of the test should be designed in accordance with the type of the test (e.g. objective tests are easy to score whereas subjective tests need rubrics to be more objective while scoring). Finally, cost of the test is also an important factor for practicality, because the test should be economical, neither cost too much or less.

### **2.4.2. Validity**

Validity is one of the valuable traits of language assessment and its existence is a must in all language tests. For a test, it is essential to be valid in order for the results to be precisely applied and explained. Validity refers to ‘accuracy of a test’ which means test should measure “what it intends to measure” (Lado, 1961:30, Hughes, 2003:26). Messick (1996) opposes the description of validity as a characteristic of a test since it is all about the



test score. In addition, Gronlund (1998:226) considers validity as “the extent to which inferences made from assessment results are appropriate, meaningful, and useful in terms of the purpose of the assessment” (in Brown, 2003:22). Chapelle (1999:254) comments on the effect of the definition of validity on language test users and she explains the reason of it as follows: “...assumptions about validity and the process of validation underlie assertions about the value of a particular type of test (e.g., "integrative," "discrete," or "performance")”.

There are four types of validity: content, criterion, construct and face validity. With respect to content validity, the test should be prepared in parallel with the goals of the subject which will be tested. The test items should represent the objectives of the test, which are aimed to be measured; therefore, it is related to the content of a test (Hughes, 2003). The second type of validity is regarded as criterion validity. The relationship between a ‘test score’ and ‘the criterion measure to be predicted’ is the criterion validity (Gronlund, 1968). In order to determine criterion validity, the criteria should be set at first. Criterion validity can be divided into two subcategories. First type of criterion validity is concurrent validity which is “established when the test and the criterion are administered at about the same time” (Hughes, 2003:27). These two tests measure the same ability and by looking at the results from each test, test administrators can determine concurrent validity. With regards to criterion validity, it is predictive validity which Black (1997:44) defines as ‘forward inference’. If a test has a predictive validity, one looking at the scores or results in a test can predict the likelihood future success of a testee (Brown, 2003). As for construct validity, testers’ interpretations from the results of a test should be in line with the theory underlying the construct that is measured (Gronlund, 1968). Therefore, this validity is the extent to which the test measures the right construct (Finocchiaro & Sako, 1983). Brown (2003) gives an example of oral interview to illustrate what construct validity is. If the theory underlying the construct of speaking ability in oral interviews includes pronunciation, fluency, grammatical accuracy, vocabulary use and sociolinguistic appropriateness and the test itself only measures pronunciation and fluency, the construct validity of that test suffers. Finally, anyone who looks at a test can comment on its face validity (Henning, 1987). Hughes (2003:33) defines it as “if it looks as if it measures what it is supposed to

measure". For instance, a speaking test that does not require testees to speak lacks face validity.

### **2.4.3. Reliability**

It is inevitable that a test instrument interferes with measurement error. Through estimating these errors, reliability is determined. In order a test to be reliable, it should include less error. Reliability can be defined as the consistency of a test or measurement (Bachman & Palmer, 1996; Brown, 2003). Reliability is about the extent to which any instrument gives the same results on repeated tryouts. Hughes (2003) points out that if the results of two tests which measure same kind of information with same people are close, the test is considered to be reliable. The more consistent the results achieved by the same participants in the same repeated measurements are, the higher the reliability of the measuring procedure will be. A test instrument, for example, can be said to be fairly reliable if a participant gets almost the same score on recurrent examinations.

In order to prepare reliable tests, there are some ways to be followed. Hughes (2003), Brown (2003) and McMillan (2007) refer to crucial factors affecting reliability. First of all, the length of a test is important, so it should be neither too long nor too short. Another factor is the reliability of scorer. The person who scores the test should be objective. Thirdly, the environmental factors play a significant role in reliability; therefore, design, equipment and acoustic of the class and the level of noise should be taken into consideration. Fourth factor is the state of test takers concerning the idea that psychological and physical situations of the test takers also interfere with the reliability. Fifth important factor is that a test should include adequate items, and the items in the test should differentiate the weak and the strong students (Hughes, 2003 and McMillan, 2007) Apart from these factors, clear test instructions and items, acquaintance of test techniques to the test takers, well-prepared scoring keys and identification of test takers by numbers instead of names have considerable impact on the reliability of the test (Hughes, 2003).

Validity exhaustively explained in section 2.4.2 and reliability mentioned above are interrelated. Chapelle (1999:255) argues that "...reliability is the prerequisite for validity". Therefore, a test which is not reliable is also not valid. Lado (1961), Henning (1987) and Hughes (2003) put emphasis on the priority of reliability over validity in constructing a test. Henning (1987:89) claims that "it is possible for a test to be reliable without being valid for a specified purpose, but it is not possible for a test to be valid without first being reliable". Though Bachman & Palmer (1996) refer to the necessity of reliability, they also point out the insufficiency of reliability alone in a test. Black (1997) also stresses that an invalid test would be out of use even if it is reliable. Despite the fact that reliability is the most crucial one among all the principles in construction of a good test, still all the principles should be considered, since none of them can be disregarded.

Küçük & Walters (2009) conducted a study in order to explore the ideas of teachers' and students' about face validity, the reliability and the predictive validity in achievement tests; and to measure the effects of face validity on predictive validity and reliability. They conducted that study with language learners and instructors at a preparatory school of a university in Turkey. The students were given two achievement tests and a final exam during a semester. In the achievement tests, they were asked questions regarding grammar, vocabulary, reading comprehension and writing. As for the final test, in addition to the skills measured in the achievement tests, speaking was tested. However, speaking was tested with a very insignificant score. Furthermore, even though listening comprehension was taught throughout the semester, it was tested in none of the tests. In addition to the tests, the students and the instructors were given two different types of questionnaires in order to investigate their perceptions about reliability and face validity. The results of the study revealed that face validity of the achievement tests mirrors both predictive validity and reliability. Besides, the examination of face validity and reliability demonstrated that there were weaknesses in testing system since the tests measured some of the language skills, not all of them. The researchers also emphasize that looking at only one aspect of language would be insufficient; therefore the tests should be examined in multiple dimensions.

#### **2.4.4. Authenticity**

Another major principle of language assessment is authenticity. A test can be called authentic if it includes real world related tasks. Stevenson (1985) describes authenticity as the requirement for testees to do a test as well as they do a daily routine. Bachman and Palmer (1996) see authenticity 'as an important test quality' (p.23). The notion 'authenticity' emerged in the 1970s when the communicative approach got on the stage and the interest increased for the 'real-life' like situations in both teaching and testing (Lewkowicz, 2000). Brown (2003) suggests several ways to make the tests more authentic: natural language should be used in the test, items should be presented in a context, topics should be meaningful, items should be thematically organized and tasks should be related to real world.

In spite of the large number of researchers' emphasis on authenticity and its' importance, there are also some scholars who oppose the idea of a test being all authentic. Raatz (1985) claims that a test cannot be authentic wholly; otherwise the test will be totally out of use. Moreover, McNamara (2000) explains the reason why he opposes a test's being more authentic than it should be that it will cost much; it will be away from simplicity and practicality.

#### **2.4.5. Washback**

Washback is a common notion in the field of language teaching and testing. The fact that testing has an influence on teaching is mentioned in education and applied linguistics literature. While 'washback' is a preferred term in British Applied Linguistics, some authors prefer the term 'backwash' (Alderson & Wall, 1993:115).

Washback is the effect of a test on both learning and teaching process (Hughes, 1989, Alderson & Wall, 1993, McNamara, 2000). For Messick (1996:241) washback is "the extent to which the introduction and use of a test influences language teachers and

learners to do things they would not otherwise do that promote or inhibit language learning".

There are two types of washback effect: positive (also called beneficial) washback and negative washback (Alderson & Wall, 1993). All assessments are thought to be prepared to have positive washback effects. Messick (1996:242) strongly emphasizes the vitality of the coherence between the activities during learning and the activities in the test for the desired beneficial washback effect. Besides the importance of the relationship between classroom and test activities, washback effect has several impacts in classroom. Spratt (2005) categorizes the washback effect's various aspects of classroom such as curriculum, materials, teaching methods, feelings, attitudes and learning (p.8).

## **2.5. Classification of Language Tests**

The classification of the language test is based on their content, frame of reference and scoring procedure, each one of which will be explained below.

### **2.5.1. Classification of Language Tests According to Their Content**

On the basis of the test content, the classification covers proficiency, achievement, aptitude, direct versus indirect, discrete-point versus integrative tests which will be explained in detail in separate parts below.

#### **2.5.1.1. Proficiency Tests**

Proficiency tests are generally administered to determine on which level a testee is and whether he or she is good enough in the subject. Harrison (1983) describes the proficiency tests as the measurement of what a learner can do with what he/she has learnt (p.8). Valette (1987) defines proficiency tests as the "global measure of ability in a language" (p.6). Brown (2003) points out the proficiency tests are not limited to only one

aspect of language, instead they measure “overall ability” (p.44). Today, TOEFL which includes listening comprehension, reading comprehension, writing, and grammar, is one of the most popular proficiency test all around the world.

### **2.5.1.2. Achievement Tests**

Achievement tests are commonly used in schools after the instruction of a unit or a subject to figure out whether the subject or the unit has been learnt by the learners and to follow their progress. Finocchiaro & Sako (1983:15) defines achievement tests are the ones “used to measure the amount and degree of control of discrete language and cultural items and of integrated language skills acquired by the student within a specific period of instruction in a specific course”. Henning (1987) describes “the probable aim of achievement tests as the certification of a language program or evaluation of the program” (p.6). Gronlund (1968) argues that the achievement tests increase ‘motivation, retention, transfer and self-understanding’ (p.3).

Linn & Gronlund (2000) categorize achievement tests as informal (teacher-made) achievement test and standardized achievement test. The former is prepared by the teachers in accordance with the subject they covered and their objectives. The latter is prepared by a committee or test publishers, considering the curriculum the teachers follow. The differences between these two tests can be discriminated in terms of learning outcomes and measurement of the content, quality of the tests, reliability, administration and scoring, and interpretation of the scores (Linn & Gronlund, 2000).

### **2.5.1.3. Language Aptitude Tests**

Aptitude tests measure the competence of a learner before s/he attends a language programme in order to predict his/her future success (Lado, 1961). They are not for measuring intelligence, but for background knowledge of a language. In other words, aptitude tests define the language level of a person. Valette (1977) describes the aptitude tests as “an indication of a person’s readiness and competence to learn the language and for

language courses, a tool to choose and level pupil according to their capabilities” (p.5). Finocchiaro & Sako (1983) put emphasis on the importance of these tests in foreseeing a person’s language learning ability and his/her probable success, besides differentiating between slow learners and fast learners. The results of these tests can be applied into the classroom when arranging classroom activities, implementing the objectives helping learners for their future plans (Linn & Gronlund, 2000).

#### **2.5.1.4. Direct versus Indirect Tests**

Hughes (2003:17) calls a test ‘direct’ “if it requires the candidate to perform precisely the skill that we wish to measure”. If a test administrator wants to measure a testee’s ability to write a composition, s/he should get that testee to write a composition. If the administrator deals with the pronunciation of a test taker, s/he should get that test taker to speak. It is highly possible to talk about a test being direct when it measures the productive skills like speaking and writing, because the ability of a testee can be observed directly. However, as for the receptive skills like listening comprehension and reading, it is essential to get a testee first to read or listen and then to show how well they have done in that process.

With respect to indirect tests, Hughes (2003:18) considers a test ‘indirect’ “if it measures the abilities that underlie the skills in which we are interested”. For instance, one section of TOEFL requires the test takers to find the inappropriate element in a sentence in order to measure their writing skill.

#### **2.5.1.5. Discrete-point versus Integrative Tests**

Discrete point tests are simple tests in which only one point of a language is tested. Valette (1987) describes discrete point tests as the measurement of a limited subject. Hughes (2003) gives the example of testing a specific grammatical structure for these tests (p.19). In contrast to discrete point tests, in integrative tests, tasks are fulfilled through the

combination of the skills and/or the sub-skills of a language. For instance, a writing test can measure spelling, vocabulary and grammar.

### **2.5.2. Classification of Language Tests According to Their Frame of Reference**

In this classification, norm-referenced and criterion-referenced tests are introduced.

#### **2.5.2.1. Norm-referenced Tests**

“Ranking” is the keyword for norm-referenced tests. A test taker is ranked in comparison to the other test takers’ achievement. While defining these tests, Montgomery & Connolly (1987) highlights the individual success in relation to the whole group success. Bond (1996) states that norm-referenced tests aim to place or award the test takers while Brown (2003) generalizes the purpose of the norm referenced tests as “to place test-takers along a mathematical continuum in rank order” (p.7). Norm-referenced tests are quantitative in terms of their results since they seek for mean, median, standard deviation and percentile rank which are statistical analysis (Klein, 1990).

#### **2.5.2.2. Criterion-referenced Tests**

In a criterion-referenced test, as mentioned in its name, the criterion has been set. Hudson & Lynch (1984) define the criterion as the cut score. Klein (1990) points out that the criterion is defined by the test items. There is a defined level for testees to be assumed successful. These tests are related with the ‘mastery and non-mastery domains’ by learners; therefore, the criterion referenced tests have qualitative results since they measure if the testees have mastered the subject or not (Klein, 1990). Typical classroom tests used at school and licensing tests are the examples of criterion-referenced tests.

### **2.5.3. Classification of Language Tests According to Their Scoring Procedure**

Objective tests and subjective tests are explained in this classification.



### **2.5.3.1. Objective Tests versus Subjective Tests**

An objective test is free from bias while scoring. These tests do not need any kind of judgment (Hughes, 2003). The correct answers do not change according to different scorers. Multiple choice tests are the common examples for objective tests. Hughes (2003) also emphasizes the popularity of these tests because of their high reliability in scoring. Coombe et al. (2007) draws attention to the scorer of these tests and state that the scorer does not need to have any special education or specific knowledge while scoring the test.

In contrast to objective test, the subjective tests need judgment in the process of evaluation (Hughes, 2003). In scoring skills such as speaking or writing, or answers to open-ended questions, sometimes the scorer's psychological or physical status, prejudice, or relation with the testee may interfere with the scoring procedure. To block, at least to lessen the interference of the scorers, rubrics are developed for subjective tests (Valette, 1987). In contrast to objective tests, subjective tests need scorers who are trained (Combee et al., 2007).

## **2.6. Standardized Testing**

Nearly all of the people who receive education are possibly exposed to standardized testing in a part of their education. As it is emphasized in its name, there are standards while improving, applying and rating these tests so that all the test takers are given tests under the same circumstances. Bagin (1989) refers to some crucial points in standardized tests such as the comparison of the students, unbiased measurement and exploration of students' capabilities. Like every test, the standardized tests have wash-back effect, either beneficial or negative. Kellaghan et al. (1982) mention the washback effects in different dimensions such as school-level, teacher-level, pupil-level and parent-level effects. These tests affect the curriculum at the school level, increase the learning of the students at the pupil level, measure the learners' current and future capabilities and success, and help teachers to define the instructions according to the needs and to choose the students for any kind of placement

at the teacher level. Bagin (1989) points out that standardized tests have beneficial washback effect since they help teachers to decide about students' success. In many countries such as the USA, Great Britain, Austria, France, Sweden, Germany, and Netherlands, standardized tests are used to measure the effectiveness of the school and the educational system. In a wide variety of countries, regional, national and international, standardized tests are regularly administered with the aim of using policy makers and making high-stake decisions instead of teacher-made tests since these tests generally tend not to be objective, reliable or valid (Riffert, 2005). In Turkey, standardized tests such as SBS (Level Identification Exam) and OSS (University Entrance Exam) are used to determine which high school or university students will attend. Today, in Turkey, 6<sup>th</sup>, 7<sup>th</sup> and 8<sup>th</sup> grade students at primary schools have to take SBS at the end of each grade in order to get into a good high school. Though the examinations are not obligatory, over 50% of students take these tests. These exams take place at the end of every academic year since it tests what students have gained in Turkish, Maths, Science, Social Science and English during the year. Each grade has English questions in their test but in different numbers. For 6<sup>th</sup> grades there are 13 questions, for 7<sup>th</sup> grades there are 15 questions and for 8<sup>th</sup> grades there are 17 questions in English section. English has the lowest coefficient, which is '1'. In these tests, such skills as listening comprehension, speaking and writing are neglected. The emphasis is mainly on grammar, vocabulary and reading comprehension.

## **2.7. Testing Language Skills**

In this section, what kind of procedure can be followed in order to measure listening comprehension, speaking, reading comprehension, writing, vocabulary and grammar is explained.

### **2.7.1. Testing Listening Comprehension**

Listening is a receptive skill which needs no production whereas it needs response. It involves understanding sounds of a language in a context. Besides, it is also seen as a way

of oral communication, and in that aspect Brown (2003) and Hughes (2003) claim that listening comprehension is a component of speaking. Buck (2001) also emphasizes the relationship between speaking and listening comprehension. He claims that in some ways listening comprehension ability is unique and in some ways it is similar to reading comprehension as they are both receptive skills.

There are many reasons to test listening skills. Every taught item should be tested to be aware of the result of a process in order to see whether it is successful or not. In other words, learners should be tested to take feedback of their learning process. “One important reason to test listening comprehension even when it might overlap quite considerably with reading comprehension is to encourage teachers to teach it” (Buck, 2001:32).

Weir (1993) categorizes listening comprehension test requirements into four as listening for direct meaning, listening for inferred meaning, listening for contributory meaning and listening for taking notes. In the first requirement, gist, main idea, details and attitude of speaker are checked. In the second one, making inferences and deductions, relating social and situational contexts, recognizing the communicative function of utterances are examined. In the third one, phonological features, grammatical notion, syntactic structure, cohesion and lexis are highlighted. In the last requirement, important points to summarize the text and selecting relevant key points are underlined.

Spoken text plays an essential role in testing listening comprehension; therefore, it demands close attention. In choosing a spoken text, Buck (2001) stresses the features that should be paid attention to, such as phonological modification (assimilation; sound influence elision; sound drop, and intrusion), accent, prosodic features (stress and intonation), speech rate, hesitations and discourse structure.

Mead & Rubin (1985) note the elements which should be included in a listening comprehension test: the listening stimuli, the questions and the test environment. The listening stimuli should include real life language. It should attract the attention and the topics should not be discriminative. In addition, the questions should not only be based on

details. Passages should include the answers of the questions. Furthermore, the testing environment should be as silent as possible. The sound quality of the system should be well-prepared and the acoustic of the room should also be taken into consideration.

In testing listening comprehension, Hughes (2003) suggests some techniques such as multiple choice, short answer, gap filling, information transfer, note taking, partial dictation, transcription, moderating the items, and presenting the texts while listening. He also strongly opposes the idea of marking grammatical or spelling errors as the aim of testing listening comprehension is to get the correct answer.

Buck (2001) categorizes listening comprehension test tasks under the approaches they belong to. In discrete-point approach, selected responses are generally used and the most frequently used tasks in this approach are phonemic discrimination tasks, paraphrase recognition tasks and lastly response evaluation tasks. In contrast to discrete-point, integrative approach examines the process of the language. Gap-filling, dictation, sentence-repetition, statement evaluation, translation are the types of integrative approach. In communicative approach, authentic texts and authentic tasks, which provide communicative purpose, are given to the learners. Rather than categorizing tests according to the approaches, Brown (2003) groups listening test tasks according to their characteristics. Intensive listening includes tasks such as recognizing phonological and morphological elements and paraphrase recognition. Responsive listening requires responses to the questions, commands, etc. Selective listening covers listening cloze tasks which require listening to fill in the gaps in given text, information transfer and sentence repetition. Extensive listening requires tasks such as dictation, communicative stimulus-response tasks.

### **2.7.2. Testing Speaking**

Though listening comprehension and speaking seem to be very much related, listening is a receptive skill whereas speaking is a productive one. Taking this important feature into account, tasks and scoring may differ in these skills.

Before preparing a test, aims, resources such as people, time, space, equipment as well as the needs and the expectations of the learner should be considered and defined well (Underhill, 1987). According to Hughes (2003), the very first thing is to specify the content while preparing a speaking test. This specification includes structures, topics, skills, type of text, rate of speech, style and accent. Enough samples should be given to guide the testees and valid sample of oral ability should be tested. Mead & Rubin (1985) suggest that there are two methods in testing speaking: the observational method and the structured method. In the observational method, the key word is 'to observe'. The tester only observes the testee with no disruption. In the structured method the tester asks the testee to perform a task for oral communication. Brown (2003) categorizes speaking test tasks such as imitative, extensive, responsive, interactive and extensive. He also thinks aural and reading comprehension cannot be separated from speaking while testing speaking. Luoma (2004) reports two kinds of speaking tasks: open-ended speaking tasks and structured speaking tasks. "Open-ended speaking tasks guide the discussion but allow room for different ways of fulfilling the task requirements. Structured speaking tasks, in contrast, specify quite precisely what the examinees should say" (p.48).

To explore the effects of task and task familiarity on oral production, Bygate & Porter (1991) conducted a study at a British University with three students who were nonnative English speakers coming from different language backgrounds. The students were interviewed at the beginning of the term and they were asked general questions about their studies, the reason why they had chosen that university and a short picture story description. After a three-month period, they were interviewed again. In the second interview, the same picture story description was asked in addition to a new one. Pauses, repairs, vocabulary and syntactic complexity were analyzed. The results of the study indicated that familiar tasks affected learners' oral performance. One student got better in fluency, the other got better in linguistic complexity and the last one got better in both of them.

Mead (1980) emphasizes the importance of interactivity, reliability and validity in scoring speaking tests. Since testing speaking is subjective, the scorer needs to prepare a rubric for scoring to be valid and reliable. Hence, the scores would be reliable and free from bias. O’Sullivan (2008) focuses on holistic and analytic scoring. Holistic scoring is simple and quick as only single mark is given, but in analytic scoring the categories which would be tested should be defined. According to The Foreign Services of Institute Scale, which is one of the most famous analytic scale, there are five categories in the process of marking in testing speaking: accent, grammar, vocabulary, fluency and comprehension. Comparing both holistic and analytic scoring, O’Sullivan (2008) suggests that there are slight differences between them.

### **2.7.3. Testing Reading Comprehension**

Even though testing reading comprehension is not as easy as it is thought, it is one of the most tested skills. There are a lot of points taken into consideration such as the aims, right text choice and right type of testing questions. There can be many choices like newspaper articles, some parts of diaries, advertisements instead of just prose. If the reader is thought not to have enough information, background knowledge should be given to make the text meaningful for the reader. Level of the text is another issue in testing reading comprehension. It should be neither easy nor complex for reader to cover.

Kitao & Kitao (1996) group reading comprehension questions according to the level of the students: Testing low level skills, testing middle or higher levels. Testing low level skills includes tasks such as word recognition, sentence recognition and matching word with a picture. In testing middle or higher levels the most popular techniques are true-false questions, multiple choice questions, short answer completion and ordering. Klinger (2004) opposes the idea of testing reading comprehension with these techniques which she finds traditional, and she thinks that traditional methods in testing reading comprehension do not look for the cognition and meta-cognition methods in reading comprehension. Instead, they just test how well a student understands a test and comments on it. She offers innovative reading comprehension assessment techniques, such as standardized norm-

referenced test, informal reading comprehension inventories, interviews and questionnaires, anecdotal records and observations, oral retelling, free-writing and think-aloud procedures to test cognitive and meta-cognitive processes.

In contrast to Kitao & Kitao (1996), Koda(2005) groups types of reading comprehension testing techniques into two: formal and informal assessment (pp.236-241). Formal assessment techniques cover free recall, in which testees define everything they recall from what they read. It is easy to prepare but difficult to score this type of assessment. Cued recall is also a kind of formal assessment. There are questions with short answers in this type. Testees do not need to recall what they read, cued recall deduces specific information. Multiple choice and cloze tests are also among the formal assessment techniques. Informal classroom assessment techniques are oral miscue analysis and observation survey. In oral miscue analysis, testees read aloud a text and tester codes the errors during reading comprehension. Observation survey aims to understand the relationship between the student and the context.

#### **2.7.4. Testing Writing**

More or less every second language learner can write in the target language but of course there are some criteria one of which can be said to be successful in writing in that language. Those criteria are determined in parallel with the aims of the program and the needs of the learners. There are a number of reasons for testing writing, and these reasons play a significant role in the aims and types of tests. Brown (2003) classifies the types of writing performance according to the aims of the test and levels of the students. Imitative writing performance can be used at a low level, intensive writing can be used for pre-intermediate level, responsive writing can be used for intermediate level, and lastly, extensive writing performance can be used for upper intermediate and advanced level students.

There are some key points to be looked for before preparing a writing test. These are the content of the test, the objectives of the test, the features of the test takers, the scoring

system and the features of the scorers, the implementation of the results, the limitations of the test validity and reliability of the test (Weigle, 2002). In addition, in the assessment process, curricular needs, expectations, realizing expectations and linguistic expectations are crucial points to be considered (Gannon, 1985). First of all, what a curriculum needs is to be thought, and besides, the teachers and the learners' needs and expectations, lack of which cause failure, are to be taken into consideration.

A study conducted by Lee & Anderson (2007) aimed to examine the validity and topic generality of writing performance test designed to place international students into appropriate ESL courses at a large mid-western university. The study was conducted with the data of graduate students' writing performance on a large-scale writing test, the ESL Placement Test at University of Illinois at Urbana-Champaign. Three factors were defined and analyzed for their interactive effect on writing performance: topical content of the writing prompts, reflection of subject background knowledge, writers' general language proficiency.

In the process of scoring the tests, it would be better not to mark spelling, punctuation and grammar errors as long as they block productivity and communication. Gannon (1985) reports two kinds of scoring: impressive and analytic. Instead of the term "impressive", "holistic" is more commonly used in the literature. Impressive scoring is the single score given with overall impression and most common scoring method among teachers because it is easy to read once instead of reading comprehension several times; therefore, it can be said to be economical. Furthermore, Weigle (2002) suggests that if large numbers of students need to be placed into writing courses with limited time and limited resources, a holistic scale may be the most appropriate choice in terms of practicality. However, it has also disadvantages. Because of single scoring, no details are given and the tester cannot diagnose the writer's ability. In contrast to the holistic one, in analytic scoring, a prepared or ready-to-use scale is used to score the piece of writing. As there is no correct answer to the writing tests because of their subjectivity, feedback is an important point to be regarded.



Weigle (2002) also refers to primary trait scoring in which products or performances are evaluated by limiting attention to a single criterion or a few selected criteria. These criteria are based upon the trait or traits determined to be essential for a successful performance on a given task. Benander et al. (2000) suggest that criteria construction in primary trait scoring takes time, and it is different from traditional scoring systems and useful form of assessment.

### **2.7.5. Testing Vocabulary**

In Merriam Websters' (2009), 'vocabulary' is defined as "a sum or stock of words employed by a language, group, individual or work or in a field of knowledge". As mentioned in the definition, vocabulary means words used in a language. In order to master a language or communicate in that language, a certain amount of vocabulary is needed.

Read (2000:2) emphasizes that "vocabulary can be seen as a priority area in language teaching, requiring tests to monitor the learners' progress in vocabulary learning and to assess how adequate their vocabulary knowledge is to meet their communication needs". There are different points of views in testing vocabulary. Some think that words can be assessed free from contexts whereas some others think tasks should be given in a context to enable interaction. In spite of the confliction, both type of vocabulary tests can be used in different situations for different purposes (Read, 2000). Lado (1961) and Hughes (2003) categorize vocabulary tests techniques into two: recognition and production techniques. Recognition technique covers multiple choice tests. In these tests, antonym, synonym or mother tongue equivalent of the word is asked and testees choose the correct answer among several options or pictures are given to fill in the blanks with the correct word. In production techniques, there are several ways such as foreign language context, picture context and native language or translation context. In the foreign language context, target language is used, which can be oral or written, to request the word which is tested. With respect to the picture one, testees are asked to look at the given picture and fill in the incomplete sentence. As for the translation one, some words are given in native language and testees are asked to translate the words into target language.

### 2.7.6. Testing Grammar

The meaning of the term “grammar” varies from context to context in which it is used. In a language context, it is the rules of accuracy in producing and sequencing phrases and sentences (Lado, 1961). Canale and Swain (1980) describe grammatical competence as the combination of the rules of phonology, the lexicon, syntax and semantics. Grammar have been popular among language teachers as Purpura (2004) emphasizes: “...language teachers have always acknowledged the inextricable link between teaching and testing and accordingly have always assessed their students’ knowledge of grammar” (p.3). Since the focus in EFL teaching is on grammar, the tests of grammar measure the ability of using the structure accurately of a target language (Madsen, 1983). Hughes (2003) highlights the place of grammatical ability the absence of which may hinder the success in skills performance; therefore, grammar takes its place in tests. Madsen (1983) categorizes grammar tests into four: tests that need limited response which include yes/no questions, true-false, pointing to something; multiple choice tests; simple completion tests such as filling in the blanks with the correct form and finally the cloze tests which require the completion of the deleted words according to the context. Apart from these types, Hughes (2003) suggests the use of paraphrasing, writing equivalent meaning of a sentence in grammar tests. Purpura (2004) proposes some tasks to make those types of grammar tests more communicative. These task types are grouped into three: Selected-response tasks, limited production tasks, extended-production tasks. First, selected-response tasks include multiple-choice and true-false activities as mentioned above, matching, discrimination, lexical list, grammaticality, noticing activities. With respect to multiple-choice tasks, they are designed to test grammatical form (morphosyntax-word order), grammatical form and meaning and identification of error. With regard to matching tasks, they are designed to measure grammatical meaning. Testees are asked to match the words and the meanings. As for discrimination tasks, they are designed to discriminate two similar grammatical knowledge areas. Testees are given pictures and asked to match the sentences with the pictures. As to noticing tasks, testees are asked to recognize some specific characteristics such as by circling the word “may” used for permission. Second, limited-production tasks aiming to test one or more areas of grammatical knowledge are gap filling, short- answer,

and dialogue (discourse) completion activities. With respect to gap-filling tasks, testees are asked to fill in the blanks with the appropriate form of the verb in order to measure grammatical form and lexical meaning. With regard to the short-answer tasks testing grammatical form and meaning, they need responses to a question, an incomplete sentence and some visuals. As for dialogue completion tasks, they ask testees to complete a dialogue in an appropriate way. Finally, in extended-production, tasks indirectly measure the grammar competence in other tests of language use such as speaking and writing. In information-gap (info-gap) tasks, (each student has information that another student in the class need to complete his/her task successfully) students are given two different information, and each student asks questions to get the information in his/her partner's card. In story-telling and reporting tasks, prompts are given to testees and they are asked to complete the story by means of their imagination. In role-play and simulations tasks, testees are given situations, aims and enough time to decide on the subject. Then, testees present their roles.

## **2.8. Types of Test Items**

Items in testing vary according to the aim of a test, type of the test, and also the skills and/or sub-skills which will be tested. There are many types of test items. In this section, multiple choice, binary choice items, gap-filling items, matching items, cloze items, paragraph writing item types will be reviewed.

### **2.8.1. Multiple Choice Items**

Multiple choice items have a question stem and choices which include the correct answer. This item type can be said to be one of the most popular test techniques. Many people prefer multiple choice items because of their easy scoring. In spite of this advantage, it is difficult to prepare multiple choice items. During preparation stage of multiple choice items, there are many points to be considered. Alderson et al. (1995) draw attention to the preparation of correct choice. There should be only one correct answer. Also the answer key

and the correct answer should be parallel. For instance if the answer key gives the correct answer as 'C', the correct answer should be 'C'. It will be better if a context is used for the multiple choice items since absence of a context may lead confusion among the choices. Another point to be considered is the appropriateness of the choices to the stem of the question. "Someone who designs houses is a ...." this stem need word choices which begin with a consonant letter (p.50).

Hughes (2003) refers to the drawbacks of this technique. In preparation stage, it is not easy to write reliable and valid items, the items "restricts what can be tested" and recognition knowledge is tested. In the administration stage, guessing and cheating may be easy. And finally negative washback may occur.

### **2.8.2. Binary Choice Items**

Yes/no or true/false items are binary choice items which have two options, one of which is correct and supposed to be not good enough since testees have the half and half chance to guess the correct answer without any knowledge (Alderson et al.,1995 & Hughes, 2003).

Some modifications can be made to improve these types of items. One of these methods is to add a third option "such as 'not given' or 'does not say'" (Alderson et al., 1995:51). The other way is to ask for the reason of the correct answer (Hughes, 2003). Besides disadvantages, this technique has several advantages. It is easy to prepare and to score the items for the administrator and also easy to answer for the testee (McMillan, 2007).

### **2.8.3 Gap Filling Items**

Testees are given sentences in which some of the words should be completed. They are supposed to fill in the blanks with appropriate word. Grammar, vocabulary, reading comprehension can be tested through these items. Completion items are easy to prepare and

the probability of guessing is less. In spite of the advantages, there are some points to be considered. Scoring of these items are not as easy as multiple choice, matching and binary choice items. The sentence should be written well, should not include clues and more than one correct answer should be avoided (McMillan, 2007). Colleagues should preview the test and a pilot test should be conducted. The list of the words to be used to fill in the gaps may be given. The instructions should be given well such as the number of the words to be used in the gaps (Alderson et al, 1995).

#### **2.8.4. Matching Items**

“Students are given a list of possible answers which they have to match with some other list of words, phrases, sentences, paragraphs or visual clues” (Alderson et al., 1995:51). This type of items includes two columns; the items on the left are premises and the items on the right are responses, and testees match the correct premise with the correct response (McMillan, 2007). McMillan (2007) points out the advantages of matching, which are the opportunity to test wide range of knowledge besides easy and objective scoring. Alderson et al. (1995) emphasize a disadvantage of this technique. For instance, if testees are given a matching test consisting of six, when five of them answered correctly, the testees do not need to think on the last item. McMillan (2007) offers a good solution to hinder such kind of a disadvantage. If the activity includes more responses than premises, the chance of guessing will be eliminated.

#### **2.8.5. Cloze Items**

In cloze tests the words are deleted no matter of what their function is. The first and the last sentence of the given passage are the same and deletion begins with the second sentence and for example every sixth word is deleted (Alderson et al., 1995). Alderson et al. (1995) suggest that the use of gap filling items is better than cloze items since in gap-filling items deletion can be determined by the person preparing the test.

### **2.8.6. Paragraph Writing**

Paragraph writing is generally used for the low level learners who are not adequate to produce long essays. A topic is given to the testees and testees are asked to write a paragraph about the topic. A variety of choices may be presented and testees are asked to choose one of the topics to write about. It is not a suggested technique since the testee will not perform on the same task (Withers, 2005). Testees should be informed about time limitation since they use time effectively. These item types are easy to prepare, but it is difficult to read and to score the answers. Scoring is subjective and can be unreliable unless a rubric is prepared for scoring (McMillan, 2007).

## **2.9 The Use of Visuals in Language Tests**

The use of visuals goes back even before the invention of written language which means visual imagery place importance in communicating and also an individual can be stimulated through the use of visuals. Benson quotes from Aristotles that, “without image, thinking is impossible” (Stokes, 2002:141) Therefore, it should take its place in language teaching and testing to trigger cognition of students. In terms of language learning, most of the learners learn best through visuals (Çakır, 2004). Pictures can be used in introduction, development and evaluation of the subject matter and also in all of the language areas of listening comprehension, speaking, reading comprehension, writing, grammar and vocabulary ( Tuttle, 1975). Interesting or entertaining pictures motivate students to respond in ways that more routine teaching aids, such as a textbook or a sentence on the board cannot (Celce-Murcia&Hilles, 1988:73, in Çakır, 2004). Hence, it should also be used in the test since the students would be more motivated and the exam will be more fun than stress. As Taylor (2006) emphasizes that the test should be made relevant and interesting through the use of visual aids such as pictures, graphics and so on ( in Ersöz, 2007:180). But the teacher should be careful in choosing right pictures for the objectives of the lesson and besides, the pictures should be seen clear especially if they are copied in black and white. Tuttle (1975) draws attention to the correct choice of the pictures otherwise it leads

misconceptions. A study has been conducted by Ginther (2001) to gain insight in the effects of using visuals in computer-based TOEFL at a Midwest university. 160 subjects participated in the study and they were grouped according to their level of English. The participants are provided different types of visuals with computer-based TOEFL multiple choice items beside questionnaires. Results of the study demonstrated that most of the participants prefer items which were accompanied by visuals.

### **2.10. English Language Teaching Curriculum in Comparison with SBS**

According to the English language teaching curriculum, the aim of teaching a foreign language in formal, informal and distance education is to help individuals gain the ability of listening comprehension, reading comprehension, speaking and writing skills and also the ability to communicate in that language and to develop positive attitudes towards foreign language teaching. Therefore, the curriculum demands teaching all of the skills in order to help students gain the overall ability for communicating in English.

After examining English language curriculum, the Regulation of Transition to Secondary Education was also examined to compare the aims of both curriculum and SBS. The aims of this regulation are below:

“SBS is conducted to identify the level of students in parallel with the gains in the curriculum. Questions in each exam may differ according to the grades. The questions are based on the syllabus of that year and the lessons appear in the weekly schedule. The exam questions are based on the gains and prepared to test interpretation, analysis, critical thinking, estimating the results and problem solving in obligatory classes such as Turkish, Mathematics, Science, Social Sciences and English” (Tebliğler Dergisi, Kasım 2007).

Coefficient of English is 1 and the number of the questions in English section differs in accordance with each grade. The number of the questions increases in parallel with the grade of the students. There are 13 questions for 6<sup>th</sup> grades, 15 questions for 7<sup>th</sup> grades and 17 questions for 8<sup>th</sup> grades in English. For 6<sup>th</sup> grade both in 2008 and 2009, there were questions covering greetings, numbers, vocabulary about family, modals such as ‘ should,

must and can', measurements 'how many, how much', simple present tense, present continuous tense, prepositions, professions, and weather. For 7<sup>th</sup> grade, in 2008, questions cover the subjects such as comparatives and superlatives, pronouns, modals (have to), simple past tense and simple perfect tense, going to future, wh- questions, illnesses, preferences, apologizing, used to and filling in the blanks according to the given picture and given sentence. In 2009, 7<sup>th</sup> grade questions cover advice(should), will, wh- questions, used to, comparatives, past tense, when clause, used to, good at, telling directions, paragraph for reading comprehension comprehension and vocabulary. In 2008, 8<sup>th</sup> grades didn't take this exam. In 2009, 8<sup>th</sup> grade exam questions cover explaining reason, agreeing, completing the dialogue according to the given speeches, present perfect tense, vocabulary on cosmetics, adjectives, preferences, simple past tense, matching the paragraph with the correct picture, daily routine, reading comprehension comprehension and vocabulary. Beside students' success in the exam, students' success average at school is also added to their overall score.

Having examined the aims of the curriculum and the Regulation of Transition to Secondary Education, the aims seem in accordance with each other. But, when the content of the English section of SBS does not seem to be parallel with the aims of not only the curriculum but also the regulation. Even though the curriculum includes listening comprehension comprehension, speaking and writing, none of the exams include these skills. The only skill tested in the exams is reading comprehension comprehension and the subskills, grammar and vocabulary.



## **CHAPTER THREE**

### **METHODOLOGY**

This chapter presents the methodological considerations and the research procedure of the study. It provides information about design of the study, data collection instruments, data collection procedures, participants of the study, piloting procedure and analysis of the collected data.

#### **3.1 Research Design**

Throughout life, people have always been in need of searching for new things. This need of change leads them to look for new ways, and tracks them to research. Research needs systematic study during the process of reaching knowledge or understanding, in this sense “research is different from other ways of knowing, such as insight, divine inspiration and acceptance of authoritative dictates” (Mertens, 1998:2)

There are several aims for research conducted in social sciences. Social research is categorized into three according to its aims: exploration, description and explanation. One of the purposes of social sciences is to explore a topic. Exploratory research generally aims at better understanding, extensive study or development of new methods for current study areas. Another purpose of social sciences is to describe events or situations. Descriptive study is used to describe the data collected in an accurate way. The last purpose of social sciences is to explain things. The explanatory research answers questions of why (Babbie, 2004).

This study is a descriptive study since it is conducted to gather information and describe the existing opinions of EFL teachers on testing and the effects of SBS on the testing process in Turkish state primary schools. Taylor (2005) defines descriptive study as description and interpretation of the present situation and explains that its aim is "... to analyze trends that are developing as well as current situations" (p.93).

Taylor and Bogdan (1984, in Brown, 1992:9) define the qualitative research as the concern of "understanding social phenomena from the actors' perspectives". Mertens (1998) states that qualitative methods are used to explore in-depth knowledge about a specific subject and refers to three factors in choosing a qualitative study. The first reason is the researchers' point of view; the second one is the nature of the research questions; and the last one is the practical reasons related to the nature of qualitative methods. Qualitative research is a suitable strategy to use particularly if there is little information about the problem (Merriam, 1998). Considering the lack of research conducted about the opinions of the EFL teachers, qualitative research is suitable for the study. Trumbull (2005) states that "qualitative researchers study things in their natural settings, attempting to make sense of, or interpret, and phenomena in terms of the meanings people bring to them" (p. 101). The researcher takes part in the qualitative research either directly or indirectly and "... qualitative research methods give real and stimulating meaning to the phenomenon" (p. 104).

Thomas (2003) highlights some points which make differences between qualitative and quantitative research methods. In qualitative methods, there is no comparison or contrast through the use of numbers. Instead, the researcher describes the characteristics of events or people, "quantitative methods, on the other hand, pay attention on measurements and amount of the characteristics displayed by people or events" (p.1). Muijs (2004) explains quantitative research methods as the numerical explanations of specific phenomena or questions and he also makes a difference between qualitative and quantitative research by arguing that the former is subjective whereas the latter is objective. According to Taylor (2005), quantitative methods aim to gather valid and objective information about facts or events. Through the use of variables, the facts or events are controlled by the researcher.

The emphasis is on the objectivity of this kind of research. Since statistical analysis is used, there is no interruption of the researcher to the result of the study.

When conducting research, the selection of data collection methods depends on the aim of the study and the topics which are explored. This study was designed as a descriptive study which includes both qualitative and quantitative data. Considering its aim and scope, this study was designed as a survey research. Mertens (1998) points out that survey is a descriptive research method which makes it possible to reach a huge population. Surveys are the ways to reach and collect data for a specific subject matter at a specific time; and the aim of surveys is to define current situations, show the relationship between the standards and the existing situation (Cohen et al. 2000). The researcher searches for the personal views to the events; and through surveys it is easy to reach for the required information since surveys can be thought as one of the best way to collect information about attitudes of the target population on a subject (Babbie, 2004). Surveys can be categorized as interviews and questionnaires (Kuş, 2003). According to Babbie (2004) typical survey is the selection of respondents and administration of a questionnaire to those respondents. In an interview, an interviewer asks the questions and respondents answer; and the interviewer records them. In a questionnaire, the questions are given to the respondents and they read the questions and write the answers on their own.

This study makes use of a questionnaire and an interview. This questionnaire included sections for both qualitative and quantitative data in order to make an in-depth description of EFL teachers' opinions on the tests they apply at school, the problems they face and the effects of SBS on their testing. The questionnaire included both open-ended and close-ended questions and semi-structured interviews were conducted as well, the data were analyzed through both qualitative and quantitative ways.

### 3.2. Participants

The participants of this study are 80 EFL teachers who are working as 6<sup>th</sup>, 7<sup>th</sup> and 8<sup>th</sup> grade English teachers at different state primary schools in Kartal, İstanbul. The participants were selected according to their accessibility by the researcher. The participants were teaching English four hours a week for each class during an academic year. Participants were selected randomly regardless of their age, gender, education status and experience. As can be seen in Table 3.1., 55 of the EFL teachers participating in the study are female and 25 of them are male.

**Table 3.1.** Gender of the Participants

Gender of Participants	Number of Participants	%
Female	55	68,8
Male	25	31,3
Total	80	100

Table 3.2. shows the university departments the participants graduated from.

**Table 3.2.** University Departments Participants Graduated from

The University Departments Participants Graduated From	Number of the Participants	%
English Language Teaching	59	73,8
English Language and Literature	11	13,8
American Culture and Literature	1	1,3
Others	9	11,3
Total	80	100

Fifty-nine of the participants were graduates of the department of English Language Teaching. Eleven of them were graduates of the department of English Language and Literature. One of them was a graduate of the department of American Culture and Literature and nine of them were graduates of various departments including physics, chemistry or philology.

Table 3.3. presents information about the minimum and the maximum work experience of the participants as EFL teachers.

**Table 3.3.** Means of the Participants' Work Experience

	<b>Minimum</b>	<b>Maximum</b>	<b>Mean</b>
<b>Work Experience</b>	3	25	9,26

The participants' work experience varied from 3 years to 25 years. All of the participants can be considered as experienced teachers since there was nobody on the first year of the teaching career. In addition, every participant can be said to be familiar with the curriculum of Ministry of Education and SBS. Moreover, each of them can be thought to have gained enough experience in preparing, administering and evaluating the examinations at school. 14 of the participants had 10 years of work experience, 11 of them had 5 years of work experience and the others' experience varied from 3 and 25 years. The mean of all participants' work experience was 9,26 years.

All of the teachers which were participating in the questionnaire taught 'My English' textbook series published by Ministry of Education. A great number of the participants were teaching all of the 6<sup>th</sup>, 7<sup>th</sup> and 8<sup>th</sup> grades; therefore, they were using all of 'My English' series.

### **3.3 Data Collection Instruments**

On the basis of the purpose and focus of the study, a questionnaire and semi-structured interviews are used to gather the information needed for the study.

#### **3.3.1 Questionnaire**

A questionnaire was developed to explore teachers' opinions on testing, the problems faced and the effect of SBS on their tests. In order to reach its final version, some

advice was taken from several lecturers. With the advice of a lecturer from Educational Sciences Department in Pamukkale University Education Faculty, the questionnaire was divided into two parts. While the first part was to collect information about the participants themselves, the second part was mainly about the data needed to answer the research questions. With respect to the content validity of the questionnaire, three lecturers from ELT department in Pamukkale University Education Faculty and a lecturer from Educational Sciences Department in Pamukkale University Education Faculty were consulted. Then, the pilot study was conducted with 12 teachers at different schools. According to their answers to the questionnaire, it took a new and its last shape.

### **3.3.1.1. Pilot Study**

The questionnaire was piloted with two sample group before conducting the actual study. In terms of content validity, twelve participants were given the questionnaire to find out if the items are clearly understood by the respondents, if there was anything confusing and accordingly make necessary changes. In terms of reliability, twenty participants answered the questionnaire.

#### **3.3.1.1.1. Piloting Procedure**

The questionnaire prepared by the researcher was made up of two parts. The first part aims to gather general information about the participants. The questions in the first part aim to look for answers concerning gender of the teacher, the university departments the participants graduated from, work experience as an English teacher, the textbooks and the grades which were instructed. The second part was designed to collect information about the research questions. This part includes questions about the opinions of EFL teachers on testing and SBS. In the second part, there are thirteen questions three of which were Likert Scale questions and the rest are open-ended questions.

The pilot group consisted of twelve English teachers all of whom were working at state primary schools and the teachers were chosen randomly according to their accessibility

and the willingness. Ten of them were female and the rest were male. Three of the teachers graduated from English Language and Literature and nine of them from English Language Teaching Department. The mean of their work experience as an English teacher was five years.

After the first pilot study some changes were made. Then the questionnaire was applied to another group consisting of twenty participants in terms of reliability. The participants were chosen randomly. Eleven of them were female and the rest were male. One of the participants graduated from American Culture and Literature, four of the participants graduated from English Language and Literature and the rest were the graduates of English Language Teaching Department. The mean of their work experience as an English teacher was 7 years.

#### **3.3.1.1.2. Results of the Pilot Study**

In the study, the questionnaire was piloted in order to avoid misunderstandings, confusion and make required renewals. According to the answers given to the questionnaire, some changes were made in the second part. The open-ended questions (from #4 to #13) were changed since the answers were not as satisfactory as expected. In parallel with the answers given, choices were added to these open-ended questions so that the answers would be much clearer, easier to answer and less time-consuming for the respondents.

As for the reliability values of second pilot study, Cronbach Alpha ( $\alpha$ ) values of two Likert Scale items in the questionnaire, the first item's Cronbach Alpha ( $\alpha$ ) value was 0.72 and the latter item's value was 0.64. Therefore, the items can be said quite reliable.

### 3.3.2. Content of the Questionnaire

The questionnaire includes 18 items and it is divided into two parts. Part-I contains questions about respondents' background information such as the departments they graduate from, age, length of work experience, grades and finally the textbooks they teach. Part-II includes questions about the opinions of the participant EFL teachers on their exams which they give at school to the grades they teach and as well as their opinions on SBS.

In part-II, some of the items are closed-format items in which the participants are given some particular choices. Some of the items demands a rating process in which the participants are asked to rate the items by using a 5-point Likert scale (ranging from "very important", 5 point to "no idea", 1 point). There are also open-ended items which require explanation from the participants. The questionnaire includes 9 questions to gather information about EFL teachers' opinions about the tests they apply.

The research questions and questionnaire items which were constructed to address each research question of the study are as follows:

#### Research Question 1

*Which language skills do state primary school EFL teachers test in their exams at school?*

In order to examine which language skills state primary school EFL teacher test in their exams at school, the first item was asked to the participants..

#### Research Question 2

*Which factors influence EFL teachers' opinions and their practices of testing?*

So as to explore the factors playing a significant role in EFL teachers' opinions and their practices of testing, items 3 to 8 in the questionnaire were constructed. Items from 3 to 8 aimed at discovering the reasons why participants test or do not test listening comprehension, speaking, reading comprehension, writing, vocabulary and/or grammar, and if they test these sub-skills/skills, what type of items they use.



### Research Question 3

*Which question types do EFL teachers prefer to use in their exams at school?*

In an attempt to discover which question types EFL teachers prefer to use in their exams at school, the second item of the questionnaire was constructed.

### Research Question 4

*To what extent do EFL teachers make use of visuals in their exams?*

The ninth item of the questionnaire was employed in order to find out to what extent EFL teachers make use of visuals in their exams since visuals are widely used in SBS.

### Research Question 5

*What are the opinions of state primary school EFL teachers on SBS?*

So as to explore what EFL teachers think about SBS, the eleventh and twelfth items in the questionnaire were used.

### Research Question 6

*What are the influences of SBS on the tests EFL teachers' apply?*

Tenth item in the questionnaire was given in order to investigate the influences of SBS on EFL teachers' testing,

Items from 10 to 13 were constructed to explore the opinions of EFL teachers considering the consistency between the curriculum and SBS and between the textbooks and SBS and also the use of supplementary materials in addition to the textbooks given by Ministry of Education for SBS.

As mentioned before, the first and the eleventh items in the questionnaire were Likert-Scale items. After the questionnaire was applied to 80 EFL teachers, Cronbach Alpha ( $\alpha$ ) values were recalculated for the main study. For the first item, the alpha reliability coefficient was 0.75, and as for the eleventh item it was 0.68. Özdamar (1999:522) gives the following values for the reliability of the questionnaires:

- $0.00 \leq \alpha < 0.40$  (The questionnaire is not reliable)
- $0.40 \leq \alpha < 0.60$  (The reliability is low)
- $0.60 \leq \alpha < 0.80$  ( The questionnaire is quite reliable)
- $0.80 \leq \alpha < 1.00$  ( The questionnaire is highly reliable)

Considering the values given and the Cronbach Alpha ( $\alpha$ ) values of the items, both the first and the eleventh item can be said to be quite reliable.

### **3.3.2 Interview**

With the goal of getting a better understanding of participants' behavior by exploring what they actually did (Silverman, 2001) through examining their experiences, opinions, feelings and knowledge of the participants (Patton, 2002); semi-structured interviews were used as a tool of qualitative research. Therefore, semi-structured interview was employed in an attempt to investigate the opinions of the participants about the exams they prepare, administer and evaluate at school besides SBS and its reflections on the exams at school. Eight teachers (corresponding to 10% of the participants having responded to the questionnaire) were interviewed in order to gain insights about a number of aspects regarding their opinions on exams and SBS.

The questions were designed by the researcher with the aim of promoting retrospective thinking on the goals of the questionnaire, whether the questionnaire helped to accomplish these goals or not. In the process of asking questions, rapport builders were taken into considerations so as to develop comfortable relationship with the respondents. In order to maintain ethics or moral qualities of the interviews, as Kvale (1996) states, the moral guidelines were followed by asking the participants' consent, using a pseudo-name for each and explaining the purpose of the interviews. Participants are called as 'interviewees' in order not to cause confusions with the participants of the questionnaire and they are given numbers from I to VIII according to the sequence of the interviews done with them (i.e. Interviewee-I).

All the interviews were scheduled between 12<sup>th</sup> and 17<sup>th</sup> April, 2010 according to the availability of the interviewees. They were individual interviews held in Turkish in teachers' room setting at schools. The interviews were tape-recorded and then translated into English by the researcher of the present study in order to determine the repeated themes. Each interview lasted approximately 15 minutes.

Table 3.4. presents information about gender of the EFL teachers who responded interview.

**Table 3.4.** Gender of the Interviewees

<b>Gender</b>	<b>Number of the Interviewees</b>
Female	5
Male	3
Total	8

As can be seen in Table 3.4., eight EFL teachers responded to interview questions. Five of the interviewees were male and three of the interviewees were female.

Table 3.5. includes information about the departments interviewees graduated from.

**Table 3.5.** University Departments Interviewees Graduated from

<b>The University Departments Interviewees Graduated from</b>	<b>N</b>
English Language Teaching	5
English Language and Literature	2
Physics Teaching	1
Total	8

Five of the interviewees were graduates of the department of English Language and Literature, two of the interviewees graduated from the English Language Teaching department and one of them was a graduate of Physics Teaching Department.

Table 3.6. presents information about the work experience of the interviewees as teachers.

**Table 3.6.** Means of the Interviewees' Work Experience

	<b>Minimum</b>	<b>Maximum</b>	<b>Mean</b>
<b>Work Experience</b>	4	10	6,1

As given in Table 3.6., interviewees' work experience varied from 4 to 10. The mean of interviewees' work experience was 6,1.

Interviewee EFL teachers were teaching 6<sup>th</sup>, 7<sup>th</sup> and 8<sup>th</sup> grade students and they were using "My English" series at the time of the interviews were conducted.

### **3.5 Data Analysis**

Since this study included both qualitative and quantitative data, the results were analyzed in different ways. For the quantitative data, the data analysis was carried out by frequency analysis. The statistical analyses were carried out using Excel software program and SPSS 12.0 software program. The results of the frequency analysis were used to interpret the results. For the qualitative data, open-ended items in the questionnaire and interviews were used.

## **CHAPTER FOUR**

### **RESULTS AND DISCUSSION OF THE FINDINGS**

#### **4.1. Introduction**

This chapter accounts for the results of the quantitative and qualitative analysis of the data collected during the procedure exhaustively explained in the previous chapter to answer the research questions of the study. The first research question aims to examine which language skills state primary school EFL teachers test in their exams. The second question seeks to explore which factors play a role in EFL teachers' opinions and practices of testing. The third question aims to discover which question types EFL teachers prefer to use in their exams at school. The fourth question seeks to find out to what extent EFL teachers make use of visuals in their exams. The fifth question aims to explore what EFL teachers think about SBS. The last question seeks to investigate the influences of SBS on EFL teachers' tests. The research questions are addressed by both a questionnaire and interviews.

#### **4.2. Results**

The results of the study are presented by following the order of these research questions as mentioned above.

#### 4.2.1 Findings on the First Research Question

The first research question in this study was:

*Which language skills/ subskills do state primary school EFL teachers test in their exams at school?*

The first research question in this study aimed to investigate which language skills state primary school EFL teachers test in their exams and the importance they give to the language skills. This research question is investigated from two perspectives: importance given to the language subskills/skills and language subskills/skills tested by the participants.

##### 4.2.1.1 Importance Given to the Language Subskills/Skills

The questionnaire results indicated that the majority of the participants placed importance on teaching reading comprehension, writing, listening comprehension, speaking, vocabulary and grammar in the process of foreign language teaching.

The percentage of the participants ranking the importance of language skills and subskills can be seen in Table 4.1.

**Table 4.1.** Percentages of Language Skills/Subskills Participants Give Importance in Their Teaching

<b>Language Subskills/skills</b>	<b>Very Important</b>	<b>Important</b>	<b>Not Important</b>	<b>Not Important at all</b>	<b>No idea</b>
Vocabulary	86%	14%	-	-	-
Speaking	62%	34%	4%	-	-
Reading comprehension	48%	46%	6%	-	-
Listening comprehension	49%	40%	10%	1%	-
Grammar	38%	56%	6%	-	-
Writing	41%	41%	15%	3%	-

Table 4.1. shows that most of the participants laid emphasis on teaching vocabulary. 86% of the participants marked teaching vocabulary as ‘very important’ while 14% of them marked

it 'important'. 62% of the participants marked teaching speaking as 'very important' and 34% of them marked it as 'important', while 4% of the participants marked it as 'not important'. 48% of the participants marked reading comprehension as 'very important', 46% of them marked it as 'important' while 6% of them marked it as 'not important'. Teaching listening was marked as 'very important' by 49%, 'important' by 40%, 'not important' by 10% and 'not important at all' by 1% of the participants. 38% of the participants marked teaching grammar as 'very important' and 56% of them marked it as 'important' while others (6%) marked teaching grammar as 'not important'. Writing was given the least importance in teaching according to the results. As for 41% of the participants teaching writing was 'very important', 41% of them found it 'important' while 15% of them found it 'not important' 3% of them found it 'not important at all'.

Table 4.2. presents the sequence of the importance given to the skills and subskills in teaching from the most to the least.

**Table 4.2.** Means of Language Skills/Subskills Participants Give Importance in Their Teaching

<b>Language Skills and Subskills</b>	<b>Mean</b>
Vocabulary	4,86
Speaking	4,59
Reading comprehension	4,41
Listening comprehension	4,36
Grammar	4,31
Writing	4,21

The results demonstrated that vocabulary that is a subskill was considered to be the most significant one among all the language skills and subskills (M=4, 86). It was followed by speaking that is a productive skill (M=4, 59), and reading that is a receptive skill (M=4, 41). Reading was followed by another receptive skill, listening, (M=4, 36) and a subskill, grammar, (M=4, 31). Another productive skill, writing, (M=4, 21) took the last place on the ranking.

Results of the interview validated the results of the questionnaire. Interviewees were asked to what extent they placed importance in teaching language skills/subskills in their classes. Besides their practices of teaching, participants were also requested to explain their reasons of teaching or not teaching the language skills/subskills. Four interviewees (corresponding to 50% of the participants interviewed) stated that they mostly taught grammar and vocabulary; and reading comprehension followed these subskills. The interviewees draw attention to SBS as an influencing factor effecting and directing their teaching practices. Language skills such as speaking and listening comprehension were disregarded.

Interviewee-III reported: *“Because of SBS, I mostly teach vocabulary and grammar. Then, I teach reading comprehension, writing and speaking. I do not teach listening comprehension since the textbooks’ listening texts are not well-recorded and clear”*.

What is more, Interviewee-V laid emphasis on inadequate time and reported: *“I do not teach listening and speaking since teaching these skills needs time and we do not have enough time. In order to meet the requirements of the curriculum concerning the textbook itself, we are in a hurry during each semester”*.

Moreover, Interviewee-VII drew attention to another issue and noted: *“In the first place, I teach reading comprehension and speaking. Grammar and vocabulary come after these skills. I do not teach listening since I could not get or download audio-CDs of the textbook. I sometimes spend time in teaching writing if I have enough time”*.

On the other hand, Interviewee-VIII with a different approach from other interviewees stressed: *“I try to keep the right balance in teaching all of the skills. The textbook includes all the skills and I do not skip any parts of the textbook”*.



#### 4.2.1.2 Language Subskills/Skills Tested

The results of the items in the questionnaire related to the issue of which language skills the participants test demonstrate that the percentage of the participants vary from one language skill to another.

Table 4.3. shows the subskills/skills and the percentages of the participants testing or not testing these subskills/skills.

**Table 4.3.** Percentage of the Participants Testing or not Testing Language Skills/Subskills

Skills /Subskills	Participants testing (%)	Participants not testing (%)
Vocabulary	100	-
Grammar	100	-
Reading comprehension	95	5
Writing	70	30
Listening comprehension	36	64
Speaking	21	79

While vocabulary and grammar were tested by all of the participants (100%), reading comprehension followed these subskills by 95%. Writing was tested by 70% of the participants. Finally, listening comprehension and speaking were tested by few of the participants (36 % and 21% respectively).

Results of the interview validated the results of the questionnaire. All of the interviewees stated that they mostly tested grammar and vocabulary. Two of the interviewees (corresponding to 25% of the participants interviewed) reported that they also tested reading comprehension. Four of them stated that they sometimes or rarely tested writing. Interviewee-I highlighted: *“I usually test grammar, vocabulary and writing. However, I do not test reading comprehension, listening or speaking”*.

Moreover, Interviewee-V and VI stated that they also tested grammar, vocabulary and sometimes writing. Interviewee-VII reported that beside grammar and vocabulary he rarely tested writing. In addition, Interviewee-II, V, VI and VII stated that they placed

importance in reading comprehension in their tests. What is more, the interviews indicated that none of the teachers prepared speaking or listening comprehension tests while some of them preferred evaluating students' classroom performances to doing such exams.

#### 4.2.2 Findings on the Second Research Question

The second research question of the study was:

*Which factors influence EFL teachers' opinions and their practices of testing?*

The second research question aimed to find out the factors influencing EFL teachers' opinions and their practices of testing. In this subsection, the results are presented in the order of the subskills/skills from the one tested most by the participants to the one tested least by those participants as specified in Table 4.3 in Section 4.2.1.2. In the questionnaire, a specific part for each subskills/skill consisted of the factors playing a role in testing these subskills/skills. The participants were asked to specify the ones influencing them.

##### 4.2.2.1 Vocabulary

As mentioned before, the results indicated that all of the participants (100%) test vocabulary. The participants were also asked the factors which influenced them to test vocabulary. Table 4.4. presents the factors influence testing vocabulary.

**Table 4.4.** Factors and Percentages of Testing Vocabulary

<b>Factors</b>	<b>%</b>
1. Because I teach vocabulary in the class.	90
2. SBS has vocabulary questions.	68
3. Curriculum covers vocabulary.	61
4. It is easy to prepare vocabulary exams.	39
5. It is easy to evaluate vocabulary exams.	33

With respect to the factors influencing the participants in testing vocabulary, the factor 'because I teach it in the class' was marked by 90 % of the participants. The second most influential factor was specified as 'SBS has vocabulary questions' by 68% of the

participants. On the third rank in terms of influence, ‘curriculum covers vocabulary’ was marked by 61% of the participants. Other factors ‘easy to prepare vocabulary exams’ and ‘easy to evaluate vocabulary exams’ were ranked on the list by 39% and 33% of the participants respectively.

The results of the interview also validated the results of the questionnaire.

In the interview, Interviewee-II reported: *“The curriculum is mainly based on vocabulary and this leads me to teach and test vocabulary. Besides, in comparison with other subskills/skills, it is easier to prepare vocabulary questions”*.

Interviewee-III, V and VI had nearly the same reasons in testing vocabulary and they put emphasis on the importance of SBS. Interviewee-V noted: *“I test vocabulary to prepare students for SBS format since the exam is based on vocabulary”*.

In addition, Interviewee-VI paid attention to another reason for testing vocabulary by stressing: *“One of the reasons of testing vocabulary is teaching vocabulary, the textbook is mainly based on vocabulary and in my exams I test what I teach”*.

Apart from the factors stated in the questionnaire, the interviews revealed other factors playing important role in testing vocabulary. For instance, Interviewee-VIII added: *“I test vocabulary because students cannot learn English without vocabulary. Furthermore, our students’ level is not good enough to test other skills, but they can succeed in vocabulary test”*.

#### **4.2.2.2 Grammar**

The results demonstrated that all of the participants (100%) test grammar. The participants were also asked the factors which influenced them to test grammar. Table 4.5. presents the factors influencing testing grammar.

**Table 4.5.** Factors and Percentages of Testing Grammar

<b>Factors</b>	<b>%</b>
1. Because I teach grammar in the class.	90
2. SBS has grammar questions.	71
3. Curriculum covers grammar.	69
4. It is easy to evaluate grammar exams.	28
5. It is easy to prepare grammar exams.	24

With regard to the factors influencing the participants in testing grammar, the factor ‘because I teach grammar’ was marked by 90% of the participants. The second most influential factor was ‘SBS has grammar questions’ was specified by 71% of the participants. On the third rank in terms of influence ‘curriculum covers grammar’ was marked by 69% of the participants. Other factors such as ‘easy to evaluate grammar exams’ and ‘easy to prepare grammar exams’ were ranked on the list by 28% and 24% of the participants respectively.

The results of the interview demonstrated that the questionnaire results were valid since the interviewees focused on almost the same reasons to test grammar in their exams.

Concerning the reason of testing grammar, Interviewee-I reported: *“First of all, the curriculum, the textbook and SBS cover grammar. Regarding the gains of the program and the syllabus, in SBS grammar is tested. I teach so that I test grammar”*.

Interviewee-II placed emphasis on the subject taught during the classes and added: *“I mainly teach grammar so that I need to test grammar to evaluate their learning process”*.

In addition, Interviewee-IV stated: *“Grammar exams are easy to prepare and evaluate. Besides, students encounter generally with grammar in SBS. As a teacher, I want my students to be successful in SBS”*.

Moreover, Participant-V drew attention to emphasis on grammar in SBS and stressed: “*SBS is mainly based on grammar. Hence, I need to teach and test grammar to prepare students for SBS*”.

#### 4.2.2.3 Reading Comprehension

The questionnaire results demonstrated that, as mentioned in Section 4.2.1, while 95% of the participants tested reading, 5% of them did not test reading in their tests at schools. The participants were also asked the factors which influenced them to test reading. Table 4.6. presents the influential factors of testing reading.

**Table 4.6.** Factors and Percentages of Testing Reading

<b>Factors</b>	<b>%</b>
Because I teach reading comprehension.	84
Reading helps to test vocabulary and grammar simultaneously.	68
Textbook includes reading skills.	62
SBS covers reading skill.	50
Curriculum covers reading.	39
It is easy to evaluate reading tests.	22
It is easy for students to answer reading tests.	17
It is easy to prepare reading tests.	16

Concerning the factors influencing the participants in testing reading comprehension, the factor ‘because I teach reading comprehension’ was marked by 84 % of the participants. The second one which 68% of the participants were influenced by was ‘reading helps us to test vocabulary and grammar simultaneously’. The third factor stated by 62% is ‘the textbook includes reading skills’. The fourth factor of 50% was that ‘SBS covers reading skill’. The fifth factor of 39% was that ‘curriculum covers reading’. Other reasons of testing reading comprehension, namely ‘it is easy to evaluate reading tests’, ‘it is easy for students to answer reading tests’ and ‘it is easy to prepare reading tests’ were marked the participants with the percentages 22%, 17% and 16% respectively.

The participants (5%) who stated that they don't test reading comprehension were also asked the factors influencing them not to test reading comprehension. Table 4.7. presents the reasons of the participants who did not test reading comprehension.

**Table 4.7.** Factors of not Testing Reading Comprehension

<b>Factors</b>
Reading comprehension is not given enough importance in SBS.
The textbook doesn't cover enough reading comprehension skills.
No reason.

The results demonstrated that 2 of the participants not testing reading skills considered the option 'reading comprehension is not given enough importance in SBS' as a reason why they did not test this skill. In addition, the reason reported by one of the participants not testing reading was 'the textbook does not cover enough reading comprehension skills'. Only one of the participants did not specify any reasons for not testing this language skill.

Results of the interview were consistent with the results of the questionnaire to some extent.

Interviewee-VI and VII reported: *"Through reading many subjects can be covered in the exam such as grammar, vocabulary and reading skills; therefore, I test reading"*.

What is more, Interviewee-V turned attention to another factor which motivated him to test reading and noted: *"I test reading skills, since the textbook covers reading skills"*.

Furthermore, Interviewee-II and VII placed emphasis on the preparation for SBS and reported: *"since reading comprehension plays an important role in SBS, I want my students to increase their familiarity with the format of SBS that's why I test reading"*.

Apart from those who test reading, Interviewee-I and III stated that they did not test reading. While Interviewee-I laid emphasis on the factor of not testing reading by stressing: *"I do not teach reading"*; Interviewee-III gave attention to the level of the students and highlighted: *The level of the students are too low to have a test of reading. I don't want them to get low marks from the exams since it is assumed to be the teacher's failure"*.

#### 4.2.2.4. Writing

The questionnaire results demonstrated that, as mentioned in Section 4.2.1, while 70% of the participants tested writing, 30% of them did not test writing in their tests at schools. The factors influencing testing writing are shown in Table 4.8. below.

**Table 4.8.** Factors and Percentages of Testing Writing

<b>Factors</b>	<b>%</b>
Writing tests measure both grammar and vocabulary simultaneously.	86
Because I teach writing.	75
Curriculum covers writing.	41
Writing tests are easy to prepare.	23

Regarding the factors having a positive influence on the ideas of teachers testing writing, ‘writing tests measure both grammar and vocabulary simultaneously’ was marked by 86% of the participants testing writing as one of those factors. The second factor for the 75% of the participants was ‘because I teach writing’. The third factor ranked on the list by 41% of them was ‘curriculum covers writing’. The last factor influencing 23% of the participants on testing writing was ‘writing tests are easy to prepare’.

Table 4.9. reflects the factors stated by the participants (30%) who did not test writing.

**Table 4.9.** Factors and Percentages of not Testing Writing

<b>Factors</b>	<b>%</b>
The textbook does not cover writing skills enough.	21
SBS does not include a writing section.	18
Evaluating writing is time-consuming.	13
I do not teach writing.	9
Other	11

The results of the questionnaire indicated that 30% of the participants did not test writing. As shown in Table 4.9., the factor ‘The textbook does not cover writing skills enough’ was marked as one of the reasons for not testing writing by 21% of the participants. ‘SBS does

not include a writing section' took the second rank on the list with the marking of 18% of the participants. The respective percentages of the participants considering 'evaluating writing is time-consuming' and 'I do not teach writing' as factors of not testing writing were 13% and 9%. In addition, 11% of the participants specified other factors in the 'other' option given below the factors in the related item of the questionnaire. The factors they mentioned were 'crowded classes', 'lack of time for teaching writing', and 'students' low level of proficiency in English'.

In accordance with the results of the questionnaire, in the interviews Interviewee-I, V and VII stated that they tested writing at times.

As for the factors influencing Interviewee-I and VII to test writing, '*because I teach writing*' is listed as number one.

Interviewee-V reported: "*I want my students to be able express themselves also in written language. Even though they cannot produce perfect sentences, I try to courage them by giving high marks. In addition, placing writing in the exam triggers the students to focus on writing, too*".

Moreover, Interviewee-VII added: "*Preparing writing exams is too easy*". On the other hand, Interviewee-VI stated that he rarely tested writing and noted: "*evaluating writing is time consuming and subjective. Sometimes I'm afraid of being unfair*".

What is more, Interviewee-II, III, IV and VIII stated that they did not test writing. With respect to the reasons of not testing, Interviewee-III and VIII highlighted that students' levels are too low to test this skill.

Furthermore, Interviewee-IV turned attention to another issue and stressed: "*evaluating writing tests is time-consuming; therefore, I do not prefer to test writing*".

Moreover, Interviewee- II stated: "*students cannot make use of time efficiently. They leave the writing part for the last minute and they cannot complete it which results in low marks even with successful students*".



#### 4.2.2.5 Listening Comprehension

According to the questionnaire results, 36% of the participants tested listening comprehension and 64 % did not test it. Table 4.10. provides information about the factors and percentages of testing listening comprehension.

**Table 4.10.** Factors and Percentages of Testing Listening Comprehension

<b>Factors</b>	<b>%</b>
I teach listening comprehension.	97
Listening comprehension is the basis for communication.	67
Curriculum covers listening comprehension.	40
We do listening activities on the net.	3

As illustrated in Table 4.10., the factor ‘I teach listening’ was marked by 97% of the participants. Other options marked as a factor were ‘listening comprehension is the basis for communication’, ‘curriculum covers listening’ and ‘we do listening activities on the net’ by the participants with percentages 67%, 40% and 3% respectively.

With regard to the participants who did not test listening comprehension, 64% of the participants stated that they did not test listening comprehension. Table 4.11. presents information about the factors and percentages of those who do not test this skill.

**Table 4.11** Factors and Percentages of not Testing Listening comprehension

<b>Factors</b>	<b>%</b>
I do not have enough equipment.	59
Students’ levels are low for testing this skill.	55
I do not have enough time to test listening comprehension.	53
I could not get audio-cds that are distributed by MOE.	45
SBS does not include a listening comprehension section.	33
It is difficult to prepare listening comprehension tests.	12
I could not download audio files from the website of MOE.	10
Testing listening comprehension is unnecessary.	4

As seen in Table 4.11., the factors leading participants not to test listening ‘I do not have enough equipment’ was marked by 59% of the participants. ‘Students’ levels are low for

testing this skill' took the second rank on the list with the marking of 55% of the participants. 'I do not have enough time to test listening' was listed on the third rank by 53% of the participants. It was followed by another factor 'I couldn't get audio-cds that are distributed by Ministry of Education' with the marking of 45% of the participants. 'SBS does not include a listening section' by 33% of the participants, 'it is difficult to prepare listening tests' by 12% of the participants, 'I couldn't download audio files from the website of Ministry of Education' by 10% of the participants and 'testing listening is unnecessary' by 4% of the participants took their places on the list of the factors leading the participants not to test listening.

As for the results of the interview, none of the interviewees tested listening skills in their exams due to the factors such as 'lack of equipment' and 'lack of time'.

Interviewee-I paid attention to the preparation and equipment by stating: *"I do not test listening due to lack of time to prepare such kind of test and lack of equipment to use in the test"*.

As for the factor affecting Interviewee- II, IV and VII, *'I don't teach listening skills so that I don't test it'* played a significant role.

Interviewee-III shared his experience and noted: *"I cannot test listening owing to the lack of equipment. I tried to test listening in my previous school but it was a total failure resulting from inadequate time and the tape recorder that did not work properly. In order to test this skill, a teacher needs a language laboratory and headphones"*.

Interviewee-VIII also emphasized the lack of proper equipment at school and also added: *"It is really hard and time-consuming to find suitable listening texts for the exam"*.

On the other hand, Interviewee-V had a different view on the issue and highlighted: *"Listening tests send successful students into a panic even when they do a listening activity in the class and I do not want to create a stressful exam atmosphere"*.

Interviewee-VI focused on students' low level of proficiency in English' and reported: *"Since some of the students are really unsuccessful and I do not want to demotivate these students with low marks from the exams"*.

#### 4.2.2.6. Speaking

The questionnaire results demonstrated that, as mentioned in Section 4.2.1, while 21% of the participants tested speaking, 79% of them did not test speaking in their exams at schools. Table 4.12. demonstrates factors influencing the participants of testing speaking.

**Table 4.12.** Factors and Percentages of Testing Speaking

<b>Factors</b>	<b>%</b>
One of the aims of learning a language is to be able to express oneself.	88
Because I teach speaking.	82
The curriculum covers speaking.	30
Communication is based on speaking.	12

The most highly marked factor motivating the participants to test speaking was 'one of the aims of learning a language is to be able to express oneself' that was specified by 88% of the participants testing speaking skill. 'Because I teach speaking' took the second rank on the list of factors marked by 82% of the participants. 'The curriculum covers speaking' and 'communication is based on speaking' were marked as other factors by the participants whose percentages were 30% and 12% respectively.

The questionnaire also revealed the reasons for the participants not to test speaking. Table 4.13. shows the factors and the percentages of not testing speaking.

**Table 4.13.** Factors and Percentages of Not Testing Speaking

<b>Factors</b>	<b>%</b>
Crowded classes.	71
I do not have enough time to test and evaluate speaking.	51
Students' low level of proficiency in English.	50
Insufficient self-esteem of the students.	41
SBS does not include a speaking section.	40
Testing and evaluating speaking tests are difficult.	35
The textbook does not include parts on speaking skills.	6
I do not teach speaking.	5
Testing speaking is unnecessary.	2

The first factor marked by 71% of the participants not testing speaking was 'crowded classes'. The second factor affecting 51% of the participants was 'I do not have enough time to test and evaluate speaking'. The third factor having influence on 50% of the participants was 'students' low level of proficiency in English'. The fourth factor influencing 41% of the participants was 'insufficient self-esteem of the students'. The fifth factor affecting 40% of them was 'SBS does not include a speaking section'. The sixth factor impressing 35% of the participants is 'testing and evaluating speaking tests are difficult'. The seventh factor affecting 6% of the participants was 'the textbook does not include parts on speaking skills'; and the eighth factor influencing 5% of them was 'I do not teach speaking'; and the last factor impressing 2% of them was that they thought 'testing speaking is unnecessary'.

The results of the interview indicated that none of the interviewees tested students' speaking skills and the factors those interviewees state were consistent with the ones marked in the questionnaire. 'Crowded classes' played the leading role in the demotivation of the interviewees to test speaking. Interviewee-III, IV, V, VII and VIII placed emphasis on the number of the students in their classrooms and time limitation. Interviewee-VII reported: *"I do not test speaking since it is impossible with these crowded classes and also we do not have enough time to spare for testing this skill"*.

Interviewee-I and VI came up with another reason apart from the ones stated above. Both of the participants focused on students' low level of proficiency in English and their

poor self-esteem. For example, Interviewee-I noted: *“I do not test speaking because students do not have self-esteem and their proficiency level is too low for this skill. They are not ready to produce”*.

In addition, Interviewee-VI reported: *“I do not want to demotivate unsuccessful students with low marks from the exams”*. Furthermore, Interviewee-II emphasizing the relationship between teaching and testing highlighted: *“I do not teach speaking; therefore, I do not test it”*.

### **4.2.3 Findings on the Third Research Question**

Third research question in this study aimed to investigate the question types EFL teachers prefer to use in their exams at school.

Third research question was:

*Which question types do EFL teachers prefer to use in their exams at school?*

This question was addressed in two ways. The first way was to look at the importance the participants placed on the question types. As for the other way, which question types were employed for which subskills/skills by the participants was examined.

#### **4.2.3.1 The Importance Given to the Question Types by the Participants**

The item related to this issue in the questionnaire employed the same kind of Likert scale as the one used in the item mentioned in Section 4.2.1. The types of questions examined in this item were as follows: multiple choice items, binary choice items, gap-filling items, matching items, open-ended questions, paragraph writing, and the last item was ‘other’ if any of the participants want to add another type of question.

Table 4.14. demonstrates the percentages of the participants ranking the importance of types of questions.

**Table 4.14.** Percentages of Question Types Participants Give Importance in Their Exams

<b>Types of Questions</b>	<b>Very Important</b>	<b>Important</b>	<b>Not Important</b>	<b>Not Important at all</b>	<b>No idea</b>
Gap-filling	41%	49%	10%	-	-
Matching	19%	64%	17%	-	-
Multiple Choice	25%	45%	25%	5%	-
Binary Choice	10%	60%	24%	6%	-
Open-ended	16%	44%	26%	14%	-
Paragraph Writing	16%	36%	29%	18%	1%

As shown in Table 4.14., the most popular question type was gap-filling. 41% of the participant found it 'very important', 49 % of them 'important and 10% of them 'not important'. Matching followed gap-filling by 19% of the participant finding it 'very important', 64% of them 'important', 17 % of them 'not important'. With respect to binary choice items 10 % of the participants stated that these types were 'very important', 60% of them 'important', 24% of them 'not important', and 6% of them 'not important at all'. As for the open-ended questions, 16% of the participants stated that it was 'very important', 44% of them 'important', 26% of them 'not important', 14% of them 'not important at all' with regard to paragraph writing, 16% of them found it 'very important, 36% of them 'important', 29% of them 'not important', 18% of them 'not important at all' and 1% of them stated that they had 'no idea'.

As for the means of the question types preferred, Table 4.15. presents the necessary information.

**Table 4.15.** Means of Question Types Preferred

<b>Types of Questions</b>	<b>Means</b>
Gap-filling	4,31
Matching	4,01
Multiple Choice	3,90
Open-ended	3,62
Binary Choice	3,50
Paragraph Writing	3,49

The results indicated that gap-filling was chosen as the most important question type (M=4,31). Matching items question type took the second place on the ranking (M=4,01). It was followed by multiple choice items (M=3,90), open-ended items (M=3,62), binary choice items (M=3,50), paragraph writing items (M=3,49).

The results of the interview validated the results of the questionnaire. Seven out of the eight interviewees stated that they used multiple choice and gap-filling items in their exams.

Interviewee-V highlighted the aim of using multiple choice items by reporting: *“I frequently use multiple choice items because I want my students to get accustomed to the format of SBS. I also use open-ended questions concerning a text. It helps me to test students’ reading comprehension comprehension and also writing ability”*.

Following these question types, matching was the third item type which was popular among the interviewees. True-false items, yes-no questions and open ended questions were also used in some of the participants’ exams.

Interviewee-VIII stated: *“gap-filling, true-false, matching and multiple choice items are popular in my exams. Sometimes, when I test reading comprehension comprehension, I also employ yes-no and open-ended questions”*. However, none of the interviewees mentioned about paragraph writing although some of them stated that they sometimes tested students’ writing skills in their exams.

#### **4.2.3.2 The Question Types Preferred to Test Subskills/Skills by the Participants**

In an attempt to investigate which test types the participants use in the process of testing each language subskills/skills, they were also asked which question types they used when they tested the subskills/skills.

Table 4.16. illustrates information about the percentages of the question types which EFL teachers preferred to test vocabulary in their exams at school.

**Table 4.16.** Question Types Preferred to Test Vocabulary and Their Percentages

<b>Question Types</b>	<b>%</b>
Matching vocabulary with their English equivalents	75
Matching pictures with the words	70
Gap-filling	70
Matching English explanation with the words	53
Matching the words with Turkish equivalents	53
Grouping the words	49
Completing the words	43

The questionnaire results indicated that 75% of the participants testing vocabulary preferred ‘matching vocabulary with their English equivalents’, 70% of them reported that they preferred ‘matching pictures with the words’. 70% of them preferred ‘gap-filling items’. 53% of them stated that they preferred items which needed ‘matching English explanation with the words’. In addition, 53% of the participants preferred items requiring the students to ‘match the words with Turkish equivalents’. 49% of the participants’ preferred ‘grouping the words’ and 43% of the participants preferred ‘completing the words’ item.

Table 4.17. gives information about the percentages of the question types which EFL teachers preferred to test grammar in their exams at school.

**Table 4.17.** Question Types Preferred to Test Grammar and Their Percentages

<b>Question Types</b>	<b>%</b>
Gap-filling items	91
Multiple choice items	85
Matching	74
True-False	59
Open-ended Questions’	53
Yes-No Questions	43

91% of the participants testing grammar preferred ‘gap-filling items’ and 85% of the participants preferred ‘multiple choice items’. The third most popular item type to test grammar was ‘matching’ employed by 74% of the participants. 59% of the participants



preferred ‘true-false items’, 53% applied to ‘open-ended questions’ and 43% of the participants chose ‘yes-no questions’ in their tests.

Table 4.18. presents information about the percentages of the question types which EFL teachers use to test reading in their exams at school.

**Table 4.18.** Question Types Preferred to Test Reading comprehension and Their Percentages

Question Types	%
Multiple choice	72
True-false	71
Yes-no questions	64
Matching	42
Open-ended questions	42

Among the participants who tested reading comprehension, 72% of them preferred ‘multiple choice questions’ in their reading tests, 71% of them preferred ‘true-false items’, 64% of them preferred ‘yes-no questions’, 42% of them preferred ‘matching items’ and 42% of them preferred ‘open-ended questions’ in their testing reading comprehension.

Table 4.19. gives information about the percentages of the question types which EFL teachers use to test writing in their exams at school.

**Table 4.19.** Question Types Preferred to Test Writing and Their Percentages

Question Types	%
Paragraph Writing	70
Rewriting a sentence using different structures	54
Describing a picture	41
Dicto-comps	18
Other	9

As for the item types the participants preferred to test writing, 70% of the participants preferred to ask students to ‘write paragraphs’, 54% of them preferred to ask students to ‘rewrite a sentence using different structures’. 41% of them asked their students to ‘describe

a picture', 18 % of them preferred dicto-comps and 9% of them preferred other techniques such as 'reordering', and 'vocabulary writing' to test writing ability of their students.

Table 4.20. illustrates the percentages of the question types which EFL teachers preferred to test listening comprehension in their exams at school.

**Table 4.20.** Question Types Preferred to Test Listening Comprehension and Their Percentages

Question Types	%
True-false	72
Multiple choice	48
Gap-filling	44
Matching	34
Short-answer questions	31
Note-taking	17
Information transfer	10

As seen in Table 4.20., with respect to the types of items employed in listening, 'true-false item' was marked by 72% of the participants. Other question types preferred by the participants were 'multiple choice' (48%), 'gap-filling' (44%), 'matching' (34%), 'short-answer questions' (31%), 'note-taking' (17%) and 'information transfer' (10%).

Table 4.21. presents information about the percentages of the question types which EFL teachers use to test speaking in their exams at school.

**Table 4.21.** Question Types Preferred to Test Speaking and Their Percentages

Question Types	%
Role-plays	76
Paired dialogues	59
Description of a picture and commenting on the picture	59
Talking on a subject	53
Open-ended questions	47
Other	6

As shown in Table 4.21., with regard to the item types employed by the participants to test speaking, 76% of the participants preferred 'role-plays', 59% of them preferred 'paired

dialogues’, 59 % of them employed ‘description of a picture and commenting on the picture’. 53% of the teachers’ preferred ‘talking on a subject’ and 47% of them employed ‘open-ended questions’ and 6% of them highlighted another technique: ‘oral summary of a story’.

#### 4.2.4 Findings on the Fourth Research Question

Due to the intense use of visuals in SBS, fourth research question in this study aimed to find out to what extent EFL teachers make use of visual in their exams and their reasons of using or not using visual.

Fourth research question was:

*To what extent do EFL teachers make use of visual in their exams?*

The results demonstrated that 86% of the participants’ used visuals and 14% of the participants did not use visuals in their tests. Table 4.22. illustrates the factors of using visuals with their percentages.

**Table 4.22.** Factors and Percentages of Using Visuals in Exams

<b>Factors</b>	<b>%</b>
Visuals help students to perceive easily.	93
Visuals make questions more clear.	85
The use of visuals makes the test more entertaining.	81
SBS includes visuals.	74

As seen in Table 4.22., the reason for the use of visuals specified by 93% of the participants was ‘visuals help students to perceive easily’. The reason marked by 85% of the participants was ‘visuals make questions more clear’. Another reason marked by 81% of them was ‘the use of visuals makes the test more entertaining’, which reduce the stress of the students’. 74% of the participants marked ‘SBS includes visuals’ as another factor leading them to use visuals.

As mentioned before 16% of the participants stated that they did not use visuals in their exam. Their reasons for not using visuals are shown with their percentages in Table 4.23.

**Table 4.23.** Factors and Percentages of not Using Visuals in Exams

Reasons	%
Visuals cannot be seen clearly in the copies.	83
It is difficult to prepare visuals for the exams.	50
Visuals take up too much space in the test paper.	33

As presented in Table 4.23., the factors demotivating the participants not to use visuals, 83% of the participants marked ‘visuals cannot be seen clearly in the copies’ as one of them. ‘It is difficult to prepare visuals for the exams’ and ‘visuals take up too much space in the test paper’ were marked as other reasons by the participants with the percentages of 50% and 33% respectively.

The results of the interview validated the results of the questionnaire. Interviewee-II, III, V, VI and VIII stated that they use visuals in their tests. Interviewee- II explained the reason: *“Since I test vocabulary, I try to make use of visuals so that the questions become more meaningful and less stressing”*. Interviewee-VI shared almost the same ideas with Interviewee-II and added: *“I use visuals in teaching vocabulary because it anchors learning. Therefore, I use visuals in my exams as it is easy for students to remember the words”*.

Interviewee- III, V and VIII emphasized the format of SBS and Interviewee-V reported: *“In SBS, visuals are widely used. In attempt to make students get accustomed to this format I use visuals in my tests”*.

Interviewee- I, IV and VII stated that they sometimes use visuals in their exams and Interviewee-I added: *“Unless I find ready-to-use exams, it is really time-consuming to prepare exams with pictures. But sometimes I prepare such kind of exams to make students get accustomed to the SBS format”*.

Interviwee-VI had a different view on the issue and highlighted: *“We don’t have good copy machines in our school. In copies the pictures are seen as only dark images and each student asks what the picture is during the exam. This situation causes noise and stress during the exam”*.

#### **4.2.5 Findings on the Fifth Research Question**

Fifth research question aimed to investigate the opinions of English teachers on SBS.

Fifth research question was:

*What are the opinions of state primary school English teachers on SBS?*

In this subsection, the results exploring the opinions of state primary school English teachers about SBS will be given through two perspectives including the consistency between the textbooks and SBS and other supplementary materials used to prepare for SBS.

##### **4.2.5.1 The Consistency between the Textbooks and SBS**

Table 4.24. presents the percentages of the opinions of participants on the sufficiency of the textbooks for SBS.

**Table 4.24.** EFL Teachers’ Opinions on the Sufficiency of the Textbooks for SBS

<b>Quite Enough</b>	<b>Enough</b>	<b>Enough but not much</b>	<b>Not Enough</b>	<b>No idea</b>
5%	22%	29%	43%	1%

As presented in Table 4.24., the questionnaire results demonstrated that the number of the participants who thought that the textbooks were not enough to prepare students for SBS (72%) outweighed the number of the participants having an idea of the sufficiency of the textbooks to prepare students (27%). As for the reasons of the inadequacy of the textbooks, several statements were produced:

Participant V stated: *“The question items used in the textbooks are not consistent with the ones used in SBS”*.

Participant XXI highlighted: *“The content of the textbook is more loaded than the one of SBS”*.

Participant XXXV added: *“Even though the textbooks cover many subskills/skills, SBS is mainly based on grammar”*.

In the interview, interviewees were also asked about their opinions on relation between SBS and curriculum, and SBS and the textbooks. The results of the interview validated the results of the questionnaire. All of the interviewees considered that not all of the gains mentioned in the regulation were tested in SBS such as writing, speaking and listening comprehension. For example, Interviewee III reported: *“I think English curriculum for primary schools should be changed from top to bottom before including English in SBS”*. In addition, Interviewee-VI noted:

*In SBS, grammar, vocabulary and reading comprehension are tested. Thus, students make efforts to be good at in these skills. Listening, speaking and writing are not tested even though the curriculum includes these skills too. There are some mistakes and deficiencies which should be recovered. If they want to test English, format should be changed. Hence, I think SBS is insufficient to test the level of the students. English should not be tested as other subjects.*

As for the sufficiency of the textbook to prepare students for SBS, except from Interviewees-I and II, all of the interviewees thought that the textbook was inadequate to prepare students and it should be improved.

Interviewee-III reported: *“I do not think the textbook is parallel with the content of SBS at least in terms of the question types. In SBS, multiple choice items are used whereas in textbook there are no multiple choice test items”*.

In addition, Interviewee-VII noted: *“The textbook is not enough for SBS preparation due to the format of the exam, which leads us to use supplementary materials”*.

Moreover, Interviewee VIII was also pessimistic about the textbooks' content but added: *"The textbook is good, but it can be improved. There might be some activities for SBS"*.

On the other hand, Interviewee-I and II shared almost the same ideas on the subject and Interviewee-I stated: *"I think SBS content and the textbook covers each other because the content of SBS is paralel with the textbook. There are no questions in SBS whose content is outside the textbook except from listening, speaking and writing skills"*.

#### 4.2.5.2 Supplementary Materials Used for SBS

Table 4.25. illustrates the percentages of supplementary materials usage of participants to prepare for SBS.

**Table 4.25.** Percentages of Participants' Supplementary Materials Usage

Participants Using Supplementary Material	76%
Participants Not Using Supplementary Material	24%

As shown in Table 4.25., the questionnaire results demonstrated that 76% of the participants employed supplementary materials to prepare their students for SBS while 24% of them did not use any supplementary materials.

Most of the participants using supplementary materials mentioned that there were several reasons why they employed this kind of materials in addition to the textbooks. The reasons of participants for using supplementary materials were as follows:

Participant-I emphasized: *"I want their students to get used to the format of SBS"*.

Participant XII added: *"The textbook is insufficient to prepare the students SBS"*.

Participant LV stated: *"I have a strong desire to have successful students according to their scores from SBS"*.

Participant LX highlighted: *“Students’ success is interpreted as teachers’ success. If the students cannot get a high score from SBS, people around school environment attribute this failure to the teacher”*.

Participant LXIV reported: *“The reason why I use supplementary materials is that I want to revise the things I do in the class and those materials give me that chance. Besides, we can get rid of the insufficiency of the textbook”*.

As for the reasons why some participants did not employ supplementary materials, the reasons below were stated:

Participant IV stated: *“I do not have enough time to do something in addition to the things that I have to go through in the class”*.

Participant XII added: *“Students cannot afford to buy supplementary materials to be used in their classes”*.

Participant LII emphasized: *“The existence of the private institutions providing courses for the students to prepare for SBS does not require the teacher at school to do extra activities, because the students have a great deal of opportunity to study through supplementary materials at these institutions”*.

#### **4.2.6 Findings on the Sixth Research Question**

The sixth research question aimed to find out what kind of influences SBS has on the tests EFL teachers apply at school.

The sixth research question was:

*What are the influences of SBS on the tests EFL teachers’ apply?*



The questionnaire results revealed that 83% of the participants were influenced by SBS in the process of preparing their tests while 17% of the participants were not affected by SBS in any way. The participants from each group came up with reasons explaining their being influenced or not by SBS.

The reasons stated by the participants influenced by SBS in their preparation of exams at their schools were as follows:

Participant III emphasized: *“I help my students to get accustomed to the format of SBS, use their time effectively and efficiently, alleviate their anxiety about the exam and get high scores”*.

Participant XVI highlighted: *“The parents expect the teacher to prepare exams in the format of SBS, thus put us under enormous pressure to take the format of SBS into consideration”*.

Participant XXIX added: *“I prepare my exams in SBS format in an attempt to increase the motivation of the students, because students have gained more positive attitudes towards the English language learning”*.

Participant LXII stated: *“It is relatively easy to evaluate the exams in the format of SBS”*.

Those who stated that they were not affected by SBS reported:

Participant VI: *“I think SBS questions are related with the syllabi of the textbooks we teach. If a teacher follows the book and tests what he teaches, I believe that students will be successful”*.

Participant XXXV: *“Many of the students prefer to go private courses and courses that enable them questions in parallel with SBS. At school, we don’t need to give SBS-like exams”*.

The results of the interview were consistent with the ones of the questionnaires. All of the interviewees admitted that they were affected by SBS while preparing their exams.

Interviewee-VI reported: *“At least in my last exams, I tried to employ multiple choice items since I want my students to get accustomed to SBS format”*.

In addition, the interviewees turned attention to a highly important issue that was the motivation aroused by something similar to SBS. Interviewee-III noted: *“SBS certainly motivates students. Several years ago when English was not tested in SBS, the students’ level of motivation concerning English as a subject was considerably low. However, today English is among the main subjects such as Turkish, Maths, Social Sciences and Science thanks to SBS”*.

Interviewee-VIII compared students’ motivation in the past and now and added: *“Students were not aware of the aims of learning a language or they were not told any aims before SBS. With the advent of SBS including questions related to English, they had an aim to learn English”*.

With regard to the reasons stated by the interviewees not influenced by SBS in preparing their tests, some interviewees stated that the classes were too crowded to use multiple-choice items in their exams because of the risk of cheating. In addition, some argued that multiple-choice technique was restrictive, thus not an accurate way to determine the proficiency level of the students in English. Besides, one interviewee suggested that the number of questions for English in SBS and their coefficient was lower than the ones of other subjects including Turkish, Math, Science and Social Sciences; therefore, it was better to focus on an attempt to enable the students to be more competent in communication in English instead of improving their recognition skills.

## **CHAPTER FIVE**

### **CONCLUSION**

#### **5.1. Introduction**

Not only does this chapter give the summary of the study with its aims and findings, but it also contains some pedagogical implications and recommendations for the further research.

#### **5.2. Overview of the Study**

This study has two main aims: (1) to explore the opinions of EFL teachers who work in state primary schools, teaching 6<sup>th</sup>, 7<sup>th</sup> and 8<sup>th</sup> graders, on the tests they give their students and (2) to discover the opinions of EFL teachers about SBS. The data collected from teachers were interpreted considering the curriculum prepared by the Ministry of Education and the syllabus of the textbook the teachers use.

Review of literature highlights that language testing passed through many stages from pre-scientific period to communicative language testing. Principles of testing, types of language tests, testing language skills and types of test items, the use of visuals in language tests are also mentioned in the related literature. In addition, the gains covered in the curriculum prepared by Ministry of Education are compared with the gains tested in the exam, namely SBS, prepared by Ministry of Education.

In this study, a questionnaire constructed by the researcher and interviews are used as data collection instruments. 80 subjects responded to the questionnaire and 8 participants were interviewed. All of the participants were EFL teachers at different state primary

schools in Kartal, Istanbul. The participants teach 6<sup>th</sup>, 7<sup>th</sup> and 8<sup>th</sup> graders, using My English series prepared by Ministry of Education. Since the questionnaire included Likert-scale, closed-format and open-ended types of items and interviews were conducted, quantitative and qualitative methods were used.

### **5.3. Conclusion**

As mentioned before, teachers at state primary schools in Turkey context have crucial roles in testing language skills/subskills. Their opinions determine the approaches and techniques to be used in testing process; therefore, their opinions gain importance. However, for the best knowledge of the researcher, there has been little research into the opinions of teachers on language testing. This study can be claimed to make contributions to the ELT field in two crucial aspects. First, the study sheds light on the opinions of EFL teachers at state primary schools on their testing practices and the factors affecting their practices. Second, the study also explores the opinions of EFL teachers on SBS. In order to get the opinions of English teachers working in state primary schools on the tests they apply and SBS, the researcher is in an attempt to reveal which skills they mostly teach and test as well as their reasons and whether they use visuals in their exams, to what extent SBS affects their tests, the consistency between curriculum and SBS, and between the textbooks and SBS.

For the first research question, the results of the questionnaire indicate that there are some contradictions between what teachers teach and test. Even though they teach listening comprehension and speaking, the least tested skills are listening comprehension and speaking. According to the results, the most popular language subskill/skill is vocabulary among the teachers since it is seen superior to all of the skills/subskills for teachers as Read (2000:2) emphasizes the place of vocabulary in teaching and testing: “vocabulary can be seen as a priority area in language teaching, requiring tests to monitor the learners’ progress in vocabulary learning and to assess how adequate their vocabulary knowledge is to meet their communication needs”. Reading and grammar are in the second place in teaching and

testing. Especially the skills including listening comprehension and speaking are the most neglected skills in testing.

For the second research question, the findings demonstrate that there are some factors influencing testing practices of EFL teachers. There are three main factors affecting on testing vocabulary and grammar: teaching in the class, SBS and curriculum. Purpura (2004) draws attention to the strong relationship between teaching and testing and teachers' needs to test grammar in parallel with what they teach. SBS, which is a standardized test, leads teachers to teach and test in accordance with it. This result is in conformity with the one of the study of Darling-Hammond & Wise (1985, in Brown, 1992:7). Results indicate that speaking is the least tested skill due to lack of time, crowded classes and also lack of confidence in students. Kitao (1996) draws attention to the difficulties in testing speaking by stating that speaking is likely to be the most difficult skill to be tested. Difficulties may be as follow: they may occur because of the pronunciation used by the speaker, the ability to get the message across, comprehension of the listener and some of them may occur because of large number of classes. Listening comprehension is also one of the least tested skills because of several reasons including lack of equipment and time, which goes hand in hand with Mead and Rubin (1985) laying emphasis on the negative or positive effects of the quality of the equipment and environment on listening tests.

For the third research question, results show that among the test item types, the most popular one is gap-filling. McMillan (2007) emphasizes that grammar, vocabulary, reading comprehension can be tested through gap-filling items and also these items are easy to prepare and the probability of guessing is less. The least popular item type is paragraph writing. MacMillan (2007) also draw attention to some discouraging points for not using paragraph writing in the tests such that these item types are easy to prepare, but difficult to read and to score the answers. Besides, scoring is subjective and it can be unreliable unless a rubric is prepared for scoring. The reasons for popular and unpopular test items stated by McMillan (2007) are in accordance with the ones given by the participants of the study.

The findings of the fourth research question indicate that the use of visuals is also popular since visuals ease perception, clarify the question and reduce stress as Taylor (2006) emphasizes that the test should be made relevant and interesting through the use of visual aids such as pictures, graphics and so on (in Ersöz, 2007:180). In spite of being popular, there are some problems with the use of visuals. Teachers encounter some problems such as unclear visuals causing chaos in the exam, difficulty in preparation of the exam and too much space in the exam paper. These motivating or discouraging effects of visuals are seen to be influential in the participants' using visual materials in either negative or positive way.

In the findings of the fifth research question, since SBS is based on vocabulary, grammar and reading comprehension, the majority of the participants think that these skills/subskills are tested to a satisfactory degree in SBS. In contrast to vocabulary, reading comprehension and grammar; other skills like writing, listening comprehension and speaking are not tested in this standardized test. Although it is defined in the Regulation of Transition to Secondary Education that SBS is conducted to identify level of students in parallel with the gains in the curriculum, the gains tested in SBS are limited to vocabulary, grammar and reading comprehension.

The majority of the participants (73%) think that the textbook is insufficient in terms of SBS. McDonough and Shaw (1993) put emphasis on the factors which lead teachers to use supplementary materials. The factors may be related with the learners, teachers, materials, methodology or examination system. The reasons stated by the participants concerning the supplementary materials are in compliance with the ones mentioned by McDonough and Shaw (1993). Nevertheless, the examination system is under the influence of SBS, thus plays the crucial role in the participants' employing these materials.

For the last research question, a great majority of the teachers (83%) confess that they are influenced by SBS in the process of constructing their own exams at schools in order to prepare students for SBS and they design at least one of their exams in parallel with SBS format as Spratt (2005) draws attention to the teaching dimension of washback effect

of a test. Since SBS is a standardized test, it has inevitable washback effects in different dimensions especially in student level and teacher-level. On the one hand, at student-level, standardized tests increase the motivation in learning. On the other hand, at teacher-level, they help teachers to define the instructions according to the needs (Kellaghan et al. 1982). The results of many studies including the one conducted by Abrams et al. (2003) indicate that teachers design their classroom assessment in parallel with the high-stake tests.

#### **5.4. Implications of the Study**

Through this study, several implications can be elicited for teachers in Turkey. The findings of this study also provide insights to those involved in the planning of language programs in terms of revising and renewing the current program in parallel with the opinions of state primary English teachers on their testing practices and SBS.

In terms of testing especially speaking and listening comprehension, the schools should be provided with necessary electronic equipment. Textbook Preparation Committees should pay more attention with the recordings of the Audio-CDs and Audio-CDs should be distributed to all of the schools in parallel with the number of the teachers instead of asking teachers to download from the website of Ministry of Education. Since some of the teachers are not good at using Internet, they may not be able to download necessary audio file for the class. Therefore; neither can they teach listening comprehension nor they can test it. In the short run, at schools, classes can be created for only English lessons, in which there should be at least visuals, tape-recorders, projectors. Through these equipment, teachers should be motivated to work efficiently and students should be motivated to learn effectively. In the long run, language laboratories should be founded in schools which include projectors, computers, headphones and microphones to ease the teaching, learning and testing process. As for the the factor ‘students’ low proficiency level in listening comprehension, speaking and also writing’ should not discourage teachers in teaching and testing these skills. Teachers should be trained in teaching and testing the so-called skills mentioned since even

the elementary level students can do listening comprehension, speaking and writing activities.

According to the results of the study, there are teachers who are not satisfied with the textbook and they turn attention to the insufficiency of activities for some skills especially for speaking, writing and reading comprehension. All English teachers should be trained in using the textbooks. The textbook and the syllabus with the curriculum gains should be introduced to all of the teachers with seminars at the very beginning of the term. At the end of the term teachers' ideas, their evaluation of the textbooks, positive and negative sides of the textbook should be discussed and each year the same gatherings both in the beginning and in the end should be repeated. The textbooks should also include some parts on SBS to prepare students for SBS format.

Standardized tests which include testing all of the skills, listening comprehension, speaking, reading comprehension, writing, grammar and vocabulary increase the ratio of emphasis on teaching listening comprehension, speaking and writing. Commissions can be constituted in Ministry of Education for preparing the ready-to-use exams in parallel with the syllabus of the textbooks. Teachers should be educated in preparing, administrating and evaluating the exams through seminars, conferences and in-service training programs. These programs should be obligatory since there are such kinds of programs in which volunteer teachers are trained.

English proficiency exams, which include all four skills and the subskills for primary schools can be conducted separate from SBS. This will trigger teachers both to teach and test and also students to learn and try hard on these skills/subskills. Therefore, language teachers will place importance to enhance the students' delivery skills, increase their confidence, and develop their methods of organization and critical thinking skills.



### **5.5. Suggestions for Further Research**

This study aimed to investigate the opinions of state primary school EFL teachers' on their testing practices at school and SBS and its effect on their tests and also the problems faced. Here are some possible ideas for further study that can be conducted on the opinions of EFL teachers on their practices of testing:

This study was carried out with only state primary school EFL teachers. Another study can be conducted with private primary school EFL teachers on their opinions and practices of testing and SBS. Thus, the results of both private and state primary EFL teachers' opinions can be compared.

Students and parents can be included in a study in order to explore their opinions about the school tests that teachers prepare, to what extent these tests are sufficient, the perceptions of students about the importance of language subskills/ skills and the effect of SBS on these tests.

## REFERENCES

- Abrams, L. M., Pedulla, J. J. & Madaus G. F. (2003). Views from the Classroom: Teachers' Opinion of Statewide Testing Programs. *Theory into Practice*, 42(1),18-29.
- Alderson, J. C. & Wall, D. (1993). Does Wasback Exist? *Applied Linguistics*, 14, 115-129.
- Alderson, J. C., Clapham, C., & Wall, D.(1995). *Language Test Construction and Evaluation*. Cambridge: Cambridge University Press.
- Babbie, E. (2004). *The Practice of Social Research*. Thomson Wadsworth:USA.
- Bachman, L. F. (1990). *Fundamental Considerations in Language Testing*. New York: Oxford University Press.
- Bachman, L. F. & Palmer, A.S. (1996). *Language Testing in Practice*. Oxford: Oxford University Press.
- Benander, R. et al (2000). Primary Trait Analysis: Anchoring Assessment in the Classroom. *The Journal of General Education*. 49(4), 279-302.
- Black, P.J. (1997). *Testing: Friend or Foe?* London: RoutledgeFalmer Press.
- Brindley, G. (2001). Assessment. In Carter, R. & Nunan, D. (Eds.). *Teaching English to speakers of other languages*. pp 137-143. Cambridge: CUP.
- Brown, D. (1992). *Altering Curricula through State Testing: Perception of Teachers and Principals*. Paper Presented at the Annual Meeting of the American Educational Research Association , San Francisco, CA, April 20-24.
- Brown, H. D. (2003). *Language Assessment: Principles and Classroom Practice*. New York: Pearson Education.
- Brown, H. D. (2001). *Teaching by principles: An interactive approach to language pedagogy*. New York: Longman.
- Buck, G. (2001). *Assessing Listening*. Cambridge: Cambridge University Press.

- Bygate, M & Porter, D. (1991). Dimension in the Acquisition of Oral Language. *ERIC*. (ED 367164).
- Canale, M. & Swain, M. (1980). Theoretical Bases of Communicative Approaches to Second Language Teaching and Testing. *Applied Linguistics*, 1, 1-47.
- Celce-Murcia, M&Hilles, S. (1988). *Techniques and Resources in Teaching Grammar*. Oxford: Oxford University Press.
- Chapelle, C. A.(1999). Validity in Language Assessment. *Annual Review of Applied Linguistics*. 19, 254-272.
- Cheng, L. (1999). Changing Assessment: Washback on Teacher Perceptions and Actions. *Teaching and Teacher Education*. 15, 253-271.
- Cimbricz, S. & College, C. (2002). State-Mandated Testing and Teachers' Beliefs and Practice. *Education Policy Analysis Archives*, 10 (2).
- Clark, J.L.D. (1983). Language Testing: Past and Current Status- Directions for the Future. *Modern Language Journal*. 67 (4), 431-443.
- Clapham, C. (2000). Assessment and Testing. *Annual Review of Applied Linguistics* 20, 147-161.
- Coombe, C. et al. (2007). *Assessing English Language Learners*. University of Michigan Press: USA.
- Cohen, L., Manion L. & Morrison K. (2000). *Research Methods in Education*. 5<sup>th</sup> ed. Routledge Falmer: London.
- Çakır, İ. (2004). Designing Activities for Young Learners in EFL Classrooms. *GÜ, Gazi Eğitim Fakültesi Dergisi*, 24 (3), 101-112.
- Darling-Hammond , L. & Wise, A. E. (1985). Beyond Standardization: State Standards and School Improvement. *Elementary School Journal*, 85(3), 315-335.
- Davies, A. (1990). *Principles of Language Testing*. Oxford: Basil Blackwell.
- Dorr-Bremme, D. W. & Herman, J. (1986). *Assesing Student Achievement: A Profile of Classroom Practices*. Center for the Study of Evaluation, University of California, Los Angeles.
- Ersöz, A. (2007). *Teaching English to Young Learners*. Ankara:EDM.
- Finocchiaro, M. & Sako, S. (1983). *Foreign Language Testing: A Practical Approach*. New York: Regents Publishing Company, Inc.

- Hamp-Lyons, L. (2000). Social, Professional and Individual Responsibility in Language Testing. *System*, 28, 579-591.
- Harrison, A. (1983). *A Language Testing Handbook*. London: Mcmillan Publishers.
- Henning, G. (1987). *A Guide to Language Testing : Development, Evaluation, Research*. Cambridge: Newbury House Publishers.
- Herman, J. Dreyfus, J. &Golan, S. (1990). The Effects of Testing on Teaching and Learning. *CSE Technical Report 327*.
- Hudson, T. & Lynch, B.(1984). A Criterion-referenced Measurement Approach to EFL Achievement Testing. *Language Testing*, 1, 171-201.
- Hughes, A. (2003). *Testing for language Teachers*. U.K.: Cambridge University Press.
- Hughes, A. (1989). *Testing for Language Teachers*. Cambridge: Cambridge University Press.
- Gannon, P. (1985). *Assessing Writing: Principles and Practice of Marking Written English*. Suffolk: Richard Clay Ltd.
- Ginther, A. (2001). *Effects of the Presence and Absence of Visuals on Performance on TOEFL-CBT Listening-Comprehension Stimuli*. Educational Testing Service Princeton, New Jersey.
- Gronlund, N. E. (1998). *Assessment of Student Achievement*. Boston: Allyn and Bacon.
- Gronlund, N. E. (1968). *Constructing Achievement Tests*. New Jersey: Prentice-Hall Inc.
- Jett, D. L. & Schafer, W. D. (1993). High School Teachers' Attitudes toward a Statewide High Stakes Student Performance Assessment. Paper Presented at the Annual Meeting of the American Educational Research Association (74<sup>th</sup> , Atlanta, GA, April 12-16) *ERIC*. (ED 357 373).
- Karavas, D.E. 1996. Using attitude scales to investigate teachers' attitudes to the communicative approach. *ELT Journal* 50 (3), 187-198.
- Kellaghan, T., Madaus, G.F., & Airasian, P.W. (1982) *The Effects of Standardized Testing*. Boston: Kluwer- Nijhoff Publishing.
- Kitao, S. K. & Kitao, K. (1996). Testing Reading comprehension. *ERIC*. (ED 398 258).
- Kitao, S. K., & Kitao, K. (1996). Testing Speaking. *ERIC*. (ED 398 261).
- Klein, T.W. (1990). Characteristics Which Differentiate Criterion-Referenced Tests from Norm-Referenced Tests. *ERIC*. (ED 324327).

- Klinger, J. K. (2004) Assessing Reading comprehension Comprehension. *Assessment for Effective Intervention*, 29 (4), 59-70.
- Koda, K. (2005). *Insights into Second Language Reading comprehension: A Cross-Linguistic Approach*. USA: Cambridge University Press.
- Kunnan, A. J. (1992). An Investigation of a Criterion-Referenced Test Using G- Theory, and Factor and Cluster Analyses. *Language Testing*, 9, 30-49.
- Kuş, E. (2003). *Nicel- Nitel Araştırma Teknikleri: Sosyal Bilimlerde Araştırma Teknikleri Nicel mi? Nitel mi?* Ankara: Anı Yayıncılık.
- Küçük, F. & Walters, J.( 2009). How good is your test? *ELT Journal*. 63 (4), 332-341.
- Kvale, S. (1996). *Interviews: An introduction to qualitative research interviewing*. Thousand Oaks, California: Sage.
- Lado, R. (1961). *Language Testing: The Construction and Use of Language Tests. A Teacher's Handbook*. Great Britain: Longman.
- Lee, H. & Anderson, C. (2007) Validity and Topic Generality of a Writing Performance Test. *Language Testing*, 24 (3),307-330.
- Lewkowitz, J. A. (2000). Authenticity in Language Testing: Some Outstanding Questions. *Language Testing*, 17 (1), 43-64.
- Linn, R L. & Gronlund N.E. (2000). *Measurement and Assessment in Teaching*. New Jersey: Prentice Hall, Inc.
- Luoma, S. (2004). *Assessing Speaking*. Cambridge: Cambridge University Press.
- Madsen, H. S.( 1983). *Techniques in Testing*. New York: Oxford University Press.
- McDonough, J.& Shaw, C. (1993). *Materials and Methods in ELT: A Teacher's Guide*. Oxford: Blackwell Publishers.
- McNamara, T. (2000). *Language Testing*. New York: Oxford University Press.
- McMillan, J., H. (2007) *Classroom Assessment: Principles and Practice for Effective Standards-Based Instruction*. U.S.A: Pearson Education Inc.
- Mead, N.A. (1980). Assessing Speaking Skills: Issues of Feasibility, Reliability, Validity and Bias. *ERIC*. (ED 197 413).
- Mead, N. A. & Rubin, D. L. (1985). Assessing Listening and Speaking Skills. *ERIC Digest*. (ED 263 626).

- Merriam, S. B. (1998). *Qualitative Research and Case Study Applications in Education*. San Francisco: Jossey-Bass Publishers.
- Mertens, D. (1998). *Research Methods in Education and Psychology: Integrating Diversity with Quantitative & Qualitative Approaches*. USA: Sage Publications.
- Messick, S. (1996). Validity and Washback in Language Testing. *Language Testing*, 13, 241-256.
- Milli Eğitim Bakanlığı Ortaöğretim Kurumlarına Geçiş Yönergesi (2007). *Tebliğler Dergisi*, Kasım 2007/ 2602
- Montgomery, P. C. & Connolly, B. H. (1987). Norm-referenced and Criterion-referenced Tests: Used in Pediatrics and Application to Task Analysis of Motor Skills. *Phys. Ther.* 67, 1873-1876.
- Muijs, D. (2004). *Doing Quantitative Research in Education with SPSS*. London: Sage Publications.
- Özdamar, K. (1999) *Paket Programları ile İstatistiksel Veri Analizi-1*, (2. Baskı), Eskişehir: Kaan Kitabevi.
- Patton, M. Q. (2002). *Qualitative research and evaluation methods*. Thousand Oaks, CA: Sage.
- Pearson, P.D., Hiebet, E.H. & Kamil, L.M. (2007) Vocabulary Assessment : What We Know and What We Need to Learn. *Reading comprehension Research Quarterly*, 42(2) 282-295.
- Popham, W. J. (2003) *Test Better, Teach Better: The Instructional Role of Assessment*. USA: [Association for Supervision & Curriculum Development](#).
- Purpura, J. E. (2004). *Assessing Grammar*. Cambridge: Cambridge University Press.
- Raatz, U. (1985). Better Theory for Better Tests. *Language Testing*, 2, 60-75.
- Read, J. (2000) *Assessing Vocabulary*. Cambridge: Cambridge University Press.
- Read, J. & Chappel, C. A. (2001) A Framework for Second Language Vocabulary Assessment. *Language Testing*, 18(1) 1-32.
- Riffert, F.(2005). The Use and Misuse of Standardized Testing: A Whiteheadian Point of View. *Interchange*, Vol. 36/1-2, 231-252.
- Shohamy, E. & Reves, T. (1985) Authentic Language Tests: Where from and Where to? *Language Testing*, 2, 48.

- Silverman, D. (2001). *Interpreting qualitative data: Methods for Analysing Talk, Text and Interaction*. London: Sage.
- Smith, L. et. al. (1986). The Role of Testing in Elementary Schools. *CSE Technical Report 321*. Arizona State University.
- Spolsky, B. (1978). Linguists and Language Testers. In Spolsky, B., editor, *Approaches to language testing*. Advances in Language Testing Series 2, Arlington, Virginia: Center for Applied Linguistics (pp. 6-11).
- Spratt, M.(2005). Washback and The Classroom: The implication for Teaching and Learning of Studies of Washback from Exams. *Language Teaching Research* 9(1): 5- 29.
- Stansfield, C. W. (2008). Where we have been and where we should go? *Language Testing*, 25(3), 311-326.
- Stevenson, D. K.(1985). Authenticity, Validity and a Tea Party. *Language Testing*, 2, 41-47.
- Stokes, S. (2002). Visual Literacy in Teaching and Learning: A Literature Perspective. *Electronic Journal of Integration of Technology in Education*, 1(1),10-19.
- Taylor, G. R. (Ed.). (2005). *Integrating quantitative and qualitative methods in research*. University Press of America.
- Taylor, S. J., & Bogdan, R. (1984). *Introduction to qualitative research methods: The search for meanings*. New York: John Wiley & Sons.
- Thomas, R. M. (2003). *Blending Qualitative and Quantitative Research Methods in Theses and Dissertations*. Thousand Oaks, California. : Corwin Press.
- Trumbull, M. (2005). *Qualitative research methods*. In Taylor, G. R. (ed.), *Integrating quantitative and qualitative methods in research*. (pp.101-126). University Press of America.
- Tuttle, H. G. (1975). Using Visual Material in the Foreign Language Classroom. *Learning Resources*, 2 (5), 9-13.
- Ulrich, L.& Turkington, C. (2001) *Get Ready! For the Standardized Tests: Reading comprehension Grade 2*. New York: McGraw-Hill Professional.
- Underhill, N. (1987). *Testing Spoken Language: A Handcourse book of Oral Testing Techniques*. Cambridge: Cambridge University Press.
- Valette, R. M. (1977). *Modern Language Testing*. USA: Harcourt Brace Jovanovich, Inc.
- Weigle, S. C. (2002). *Assessing Writing*. UK: Cambridge University Press.

Weir, C. J. (1990). *Communicative Language Testing*. New York: Prentice Hall.

Weir, C. (1993). *Understanding and Developing Language Tests*. New York: Prentice Hall.

Williams, M., & Burden, R.L. (1997). *Psychology for Language Teachers: A Social Constructivist Approach*. Cambridge: Cambridge University Press.

Wright, T. (1987). *Roles of Teachers and Learners*. Oxford: Oxford University Press.

## INTERNET SOURCES

Bagin, Carolyn Boccella, (1989). Talking to Your Child's Teacher about Standardized Tests. *Practical Assessment, Research & Evaluation*, 1(5). Retrieved January 22, 2010 from <http://PAREonline.net/getvn.asp?v=1&n=5>

Bond, Linda A. (1996). Norm- and criterion-referenced testing. *Practical Assessment, Research & Evaluation*, 5(2). Retrieved February 2, 2010 from <http://PAREonline.net/getvn.asp?v=5&n=2>

Merram-Websters' Dictionary. Retrieved March 30, 2009, from [http://www.merriam-webster.com/dictionary/vocabulary.\(2a\)](http://www.merriam-webster.com/dictionary/vocabulary.(2a))

Mevzuat Bankası, Milli Eğitim Bakanlığı Yabancı Dil Eğitimi ve Öğretimi Yönetmeliği, İkinci Kısım, Madde 5. Retrieved April 10, 2009 from <http://www.meb.gov.tr>

O'Sullivan, B. (2008). Notes on Assessing Speaking. Retrieved February 23, 2009, from [www.lrc.cornell.edu/events/past/2008-2009/papers08/osull1.pdf](http://www.lrc.cornell.edu/events/past/2008-2009/papers08/osull1.pdf)

Stokes, S. Visual Literacy in Teaching and Learning: A Literature Perspective. Retrieved March 28, 2010 from *Electronic Journal for the Integration of Technology in Education*, vol. 1, no. 1. [ejite.isu.edu/Volume1No1/pdfs/stokes.pdf](http://ejite.isu.edu/Volume1No1/pdfs/stokes.pdf)

Taylor, L (2006). *Young Learner Testing*. Retrieved July 2006, from [http:// www.cambridgeesol.org/teach/yle/](http://www.cambridgeesol.org/teach/yle/)

Withers, G. (2005). Item Writing for Tests and Examinations. Retrieved March 16, 2010 from <http://www.unesco.org/iiep>.

Wharton, Sue. Characterising Your Language Test- Language Testing. Retrieved May 28, 2009, from <http://www.philselfsupport.com/testing1.htm>

What is a standardized test? Retrieved January 22, 2010, from [http://www.pearsonedmeasurement.com/research/faq\\_1a.htm](http://www.pearsonedmeasurement.com/research/faq_1a.htm)



## APPENDICES

Appendix 1  
Questionnaire

Değerli Meslektaşım,

**“Devlet ilköğretim okullarındaki 6, 7 ve 8. sınıflara uygulanan İngilizce sınavları, SBS’nin bu sınavlar üzerindeki etkileri ve karşılaşılan sorunlar hakkında İngilizce öğretmenlerinin düşünceleri”**ni konu alan bir tez araştırması yapmaktayım. İlgili bilgiye ulaşabilmem için sizin cevaplarınız, düşünceleriniz ve tecrübeleriniz büyük önem taşımaktadır. Bu yüzden zamanınızın bir kısmını ayırarak aşağıdaki soruları cevaplamanızı rica ediyorum. Bütün cevaplar kesinlikle sadece araştırma amaçlarına yönelik olarak kullanılacaktır.

Teşekkürler.

Tuğba AKINCI

Pamukkale Üniversitesi İngiliz Dili Eğitimi  
Yüksek Lisans Programı Öğrencisi

**BÖLÜM I**

**1- Cinsiyetiniz**     Bay                       Bayan

**2- Mezun olduğunuz bölümü lütfen işaretleyiniz.**

- İngiliz Dili Eğitimi (İngilizce öğretmenliği)  
 İngiliz Dili ve Edebiyatı  
 Amerikan Dili ve Edebiyatı  
 Diğer (Belirtiniz).....

**3- Kaç yıldır İngilizce öğretmeni olarak görev yapmaktasınız?**

.....



**3- Sınavlarınızda dinleme becerisinin ölçülmesine yer veriyor musunuz?**

- EVET,çünkü** (Birden fazla seçeneği işaretleyebilirsiniz)  
 Derste yer veriyorum  
 Müfredatta yer verilmiş.  
 İletişim için temel öge duyduğunu anlamadır.  
 Öğrencilerime internetten çalışmalar yaptırıyorum.  
 Diğer(Belirtiniz).....

- HAYIR,çünkü** (Birden fazla seçeneği işaretleyebilirsiniz)  
 Dinleme sınavı hazırlamak zor.  
 Yeterli donanım yok.  
 Yeterli zaman yok.  
 Öğrenci seviyesi yetersiz.  
 SBS' de yer verilmiyor.  
 Gereksiz olduğunu düşünüyorum.  
 CD elime ulaşmadı.  
 MEB web sitesinden dinleme ses dosyalarını indiremedim.  
 Diğer(Belirtiniz).....

**3. soruya cevabınız 'EVET' ise, bu beceriyi ölçmek için aşağıdaki soru tiplerinden hangisini kullanıyorsunuz?**

- Çoktan seçmeli  
 Doğru yanlış  
 Boşluk doldurma  
 Eşleştirme  
 Kısa cevaplı sorular  
 Bilgi transferi (information transfer)  
 Not alma (note taking)  
 Diğer(Belirtiniz).....

**4- Sınavlarınızda konuşma becerisinin ölçülmesine yer veriyor musunuz?**

- EVET, çünkü** (Birden fazla seçeneği işaretleyebilirsiniz)  
 Derste yer veriyorum  
 Müfredatta yer veriliyor  
 Dil öğrenmenin amaçlarından biri kendini ifade edebilmedir.  
 Diğer(Belirtiniz).....

- HAYIR, çünkü** (Birden fazla seçeneği işaretleyebilirsiniz)  
 Sınıf mevcutları kalabalık  
 Ölçmek ve değerlendirmek için yeterli zaman yok  
 Kitapta yer verilmemiş  
 Derste yer vermiyorum  
 Bu tür sınavı hazırlamak zaman alıcı  
 Ölçmek ve değerlendirmek zor

- Gereksiz olduğunu düşünüyorum
- SBS' de yer verilmiyor
- Öğrencilerin İngilizceyi kullanmaya yönelik özgüveni yetersiz
- Öğrencilerin seviyesi yetersiz.
- Diğer(Belirtiniz).....

**4. soruya cevabınız 'EVET' ise konuşma becerisini nasıl ölçtüğünüzü belirtiniz**

- Bir konu üzerinde konuşurma
- Verilen resim hakkında betimleme ve yorum yaptırma
- Açık uçlu sorular sorma
- Rol yaptırma (Role play)
- Tartışma ve- veya karşılıklı konuşurma
- Diğer(Belirtiniz).....

**5- Sınavlarınızda öğrencinin okuma becerisini ölçen sorulara yer veriyor musunuz?**

- EVET, çünkü** (Birden fazla seçeneği işaretleyebilirsiniz)
- Derste yer veriyorum.
- Kitapta yer veriliyor.
- Müfredatta yer veriliyor.
- SBS' de yer veriliyor.
- Hazırlamak kolay.
- Değerlendirmek kolay.
- Kelime ve dilbilgisini aynı anda ölçmeyi sağlıyor.
- Öğrenciler cevaplamakta zorlanmıyor.
- Diğer(Belirtiniz).....

- HAYIR, çünkü** (Birden fazla seçeneği işaretleyebilirsiniz)
- Derste yer vermiyorum.
- Kitapta yeteri kadar yer verilmiyor .
- Hazırlamak zaman alıyor.
- SBS' de gereken önem verilmiyor.
- Diğer(Belirtiniz).....

**5. soruya cevabınız 'EVET' ise bu beceriyi ölçmek için hangi soru tiplerini kullandığınızı belirtiniz.**

- Çoktan Seçmeli
- Eşleştirme
- Evet- Hayır soruları
- Doğru yanlış
- Açık uçlu sorular
- Diğer (Belirtiniz).....

**6- Sınavlarınızda yazma becerisinin ölçülmesine yer veriyor musunuz?**

- EVET, çünkü** (Birden fazla seçeneği işaretleyebilirsiniz)
- Derste yer veriyorum.
- Müfredatta yer veriliyor.
- Hazırlamak kolay.
- Kelime ve dilbilgisini aynı anda ölçmeyi sağlıyor.
- Diğer (Belirtiniz).....

- HAYIR, çünkü** (Birden fazla seçeneği işaretleyebilirsiniz)
- Derste yer vermiyorum.
- Kitapta yeteri kadar yer verilmiyor.
- Değerlendirmek zaman alıyor.
- SBS' de yer verilmiyor.
- Diğer (Belirtiniz).....

**6. soruya cevabınız 'EVET' ise yazma becerisini nasıl ölçtüğünüzü belirtiniz**

- Paragraf yazdırma
- Duyduğunu kendi ifadelerini kullanarak yazdırma (Dicto-comp)
- Verilen resmi betimletme
- Verilen cümleyi farklı yapı kullanarak yeniden yazdırma
- Diğer (Belirtiniz).....

**7- Sınavlarınızda derste öğrettiğiniz sözcüklerin ölçülmesine yer veriyor musunuz?**

- EVET, çünkü** (Birden fazla seçeneği işaretleyebilirsiniz)
- Derste yer veriyorum
- Müfredatta yer veriliyor
- SBS' de yer veriliyor
- Hazırlamak kolay
- Değerlendirmek kolay
- Diğer(belirtiniz).....

- HAYIR, çünkü** (Birden fazla seçeneği işaretleyebilirsiniz)
- Derste yer vermiyorum
- Hazırlamak zaman alıyor
- Diğer(belirtiniz).....

**7. soruya cevabınız 'EVET' ise, aşağıdaki soru tiplerinden hangisini kullanıyorsunuz?**

- Sözcüğü İngilizce anlamıyla eşleştirme
- Sözcük- resim eşleştirme
- Sözcük tamamlama
- Sözcüğün Türkçe karşılığını yazma
- Sözcüğün İngilizce karşılığını yazma
- Sözcük gruplama
- Sözcüğü cümle içinde kullanma

Diğer(belirtiniz).....

**8- Sınavlarınızda dilbilgisinin ölçülmesine yer veriyor musunuz?**

- EVET, çünkü** (Birden fazla seçeneği işaretleyebilirsiniz)
- Derste yer veriyorum.
- Müfredatta yer veriliyor.
- SBS' de yer veriliyor.
- Hazırlamak kolay.
- Ölçmek ve değerlendirmek kolay.
- Diğer(belirtiniz).....

**HAYIR, çünkü**

.....

.....

.....

**8.soruya cevabınız EVET ise kullandığınız soru tipleri hangisi ya da hangileridir? (Birden fazla seçeneği işaretleyebilirsiniz)**

- Çoktan Seçmeli
- Eşleştirme
- Boşluk doldurma
- Doğru yanlış
- Açık uçlu sorular
- Evet- hayır soruları
- Diğer (Belirtiniz).....

**9- Sınavlarınızda görsel malzeme ( resim, şekil, şema, tablo, vb.) kullanıyor musunuz?**

- EVET, çünkü** (Birden fazla seçeneği işaretleyebilirsiniz)
- Algıda kolaylık sağlıyor.
- Soruların anlaşılır olmasını sağlıyor.
- Sınavı eğlenceli hale getiriyor.
- SBS'de kullanılıyor .
- Diğer (Belirtiniz).....

- HAYIR, çünkü** (Birden fazla seçeneği işaretleyebilirsiniz)
- Hazırlamak zor oluyor.
- Çok yer kaplıyor.
- Fotokopide görsel malzeme net görünmüyor.

Diğer (Belirtiniz).....

**10- SBS' de 2008 yılından bu yana İngilizce soruları da yer almaktadır. Bu durum sınavlarınızı SBS' ye yönelik hazırlamanızda sizi etkiliyor mu?**

**EVET,çünkü** .....

**HAYIR, çünkü** .....

**11- Sizce SBS' de sorulan soru tipleri Milli Eğitim Bakanlığı İngilizce müfredatında bahsi geçen kazanımları ne derecede ölçüyor?**

	Çok Yeterli (5)	Yeterli (4)	Az Yeterli (3)	Yetersiz (2)	Bir Fikrim Yok (1)
a) dilbilgisi	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
b) kelime	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
c) okuma	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
d) yazma	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
e) dinleme	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
f) konuşma	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

**12- Kullandığınız ders kitabı tek başına öğrencileri SBS' ye hazırlamada sizce ne kadar yeterli?**

Çok Yeterli  Yeterli  Az Yeterli  Yetersiz  Bir Fikrim Yok

**Düşüncenizi belirtiniz**

.....  
 .....  
 .....

**13- SBS'ye hazırlık için ek kaynak kullanıyor musunuz? Lütfen cevabınızı gerekçesiyle belirtiniz.**

**EVET, çünkü** .....

**HAYIR, çünkü** .....

---

---

**TEŞEKKÜRLER**

## Appendix 2

## Interview Questions

- 1- Which language skills/subskills do you mostly teach?
- 2- Which language skills/subskills do you mostly test? Please explain your reasons.
- 3- Which question types do you mostly use in your tests?
- 4- Do you make use use of visuals in your tests? Please explain your reasons.
- 5- What do you think about the relationship between SBS and curriculum, SBS and the textbooks?
- 6- Does SBS affect your tests?



## Appendix 3

## Reliability Outputs of Pilot Study

**Item 1****Reliability Statistics**

Cronbach's Alpha	N of Items
,722	6

**Item 11****Reliability Statistics**

Cronbach's Alpha	N of Items
,647	6

## Appendix 4

## Reliability Outputs of Main Study

**Item 1****Reliability Statistics**

Cronbach's Alpha	N of Items
,759	6

**Item 11****Reliability Statistics**

Cronbach's Alpha	N of Items
,682	6

Appendix 5

2009 SBS Questions for 6<sup>th</sup>, 7<sup>th</sup> and 8<sup>th</sup> Grades

6. SINIF

YABANCI DİL TESTİ

A

BU GRUPTAKİ SORULAR YABANCI  
DİLLERE AİTTİR. OKUDUĞUNUZ YABANCI  
DİLE AİT SORULARI CEVAPLANDIRINIZ.

## İNGİLİZCE

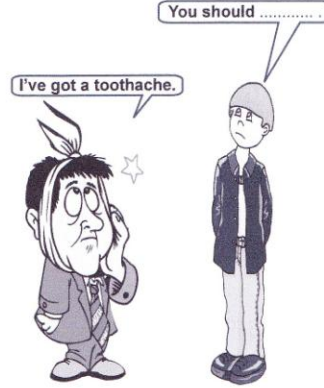
1.- 10. sorularda boş bırakılan yerlere uygun  
düşen sözcük veya ifadeyi bulunuz.

1.



- A) Taking photographs is my hobby.  
B) I'm good at basketball.  
C) I enjoy cycling.  
D) We love dancing.

2.



- A) wash your hands often  
B) see a dentist  
C) drink hot chocolate  
D) help the doctor

3. Emre'nin kilosu ile ilgili  
söyleyebileceği seçeneği  
işaretleyiniz.



- A) give up eating junk food  
B) do my homework  
C) eat more  
D) play computer games

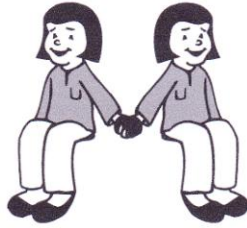
## 6. SINIF

## YABANCI DİL TESTİ

A

4.

Özge is my .....  
We're twins.



Özge

Gözde

- A) father  
B) mother  
C) uncle  
D) sister

5.



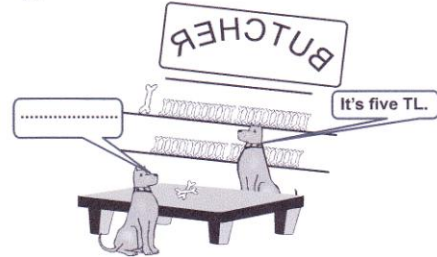
- A) Call the doctor!  
B) Don't panic!  
C) Keep your room clean!  
D) Calm down!

6.



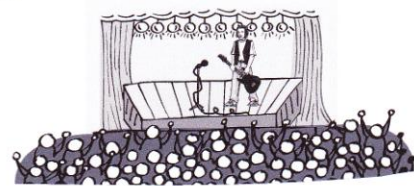
- A) a lot of sugar  
B) an umbrella  
C) a radio  
D) some newspapers

7.



- A) How many do you want?  
B) What's your favourite toy?  
C) How much is that big bone?  
D) What's your job?

8.



He's going to .....

- A) play the guitar.  
B) make a sandwich.  
C) watch television.  
D) break the microphone.

## 6. SINIF

## YABANCI DİL TESTİ

A

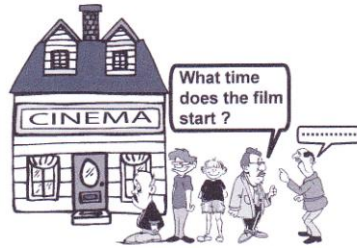
9.



They are .....

- A) sleeping                      B) camping  
C) studying                      D) running

10.



- A) At eleven o'clock.  
B) Tomorrow night.  
C) On Monday.  
D) Yesterday morning.

11.



Okan Sevgi Murat Barış İbrahim

- A) Sevgi                              B) Okan  
C) Barış                              D) Murat

12.

I'm listening to my teacher.  
She's teaching Geography now.  
Where am I?



- A) On the school-bus.  
B) In the canteen.  
C) In the classroom.  
D) At the laboratory.

13.

I'm small and green.  
I've got four legs. I live in lakes.  
I can swim and jump but I can't fly.  
I like flies and lettuce. Who am I?

?



7. SINIF

YABANCI DİL TESTİ

A

BU GRUPTAKİ SORULAR YABANCI DİLLERE AİTTİR. OKUDUĞUNUZ YABANCI DİLE AİT SORULARI CEVAPLANDIRINIZ.

## İNGİLİZCE

1.- 10. sorularda boş bırakılan yerlere uygun düşen sözcük veya ifadeyi bulunuz.

1.



- A) do your homework.  
B) take your sunglasses.  
C) wash your hands.  
D) wear your coat.

2.



- A) He's positive.  
B) You'll look beautiful in white.  
C) It smells delicious.  
D) I mean it's fashionable.

3.



- A) prefers shoes.  
B) isn't good at fishing.  
C) doesn't like fish.  
D) is interested in shoes.

4.



- A) Who B) How C) Where D) Which

5.



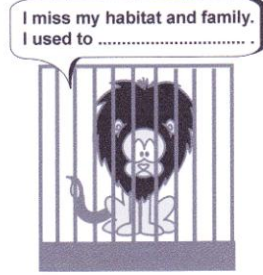
- A) worse B) better  
C) more important D) more difficult

## 7. SINIF

## YABANCI DİL TESTİ

A

6.



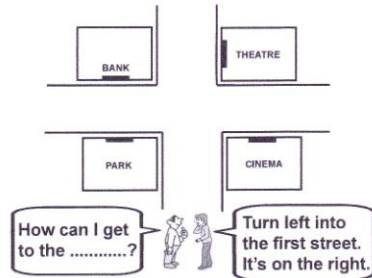
- A) drive a car
- B) wear glasses
- C) be unhappy then
- D) live in the jungle

7.



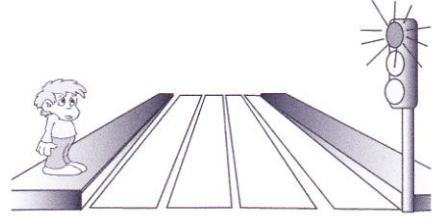
- A) I was a baby.
- B) we were at the university.
- C) we were in the classroom.
- D) I was a student.

8.



- A) theatre
- B) park
- C) bank
- D) cinema

9. The traffic light is red. He must .....



- A) cross
- B) wait
- C) not stop
- D) park his car

10. Resmi dikkate alarak falcının söylediği ifadeyi işaretleyiniz.



- A) You will have an accident soon.
- B) You will feel quite good today.
- C) It will be a good week for you.
- D) Don't worry! Everything will be okay.



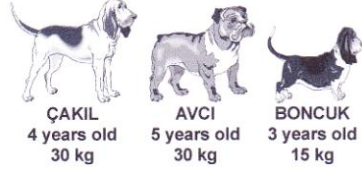
## 7. SINIF

## YABANCI DİL TESTİ

A

11 - 12. sorularda verilen resme göre doğru olan seçeneği işaretleyiniz.

11.



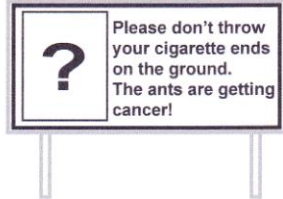
- A) Boncuk is older than Avcı.  
B) Çakıl is younger than Boncuk.  
C) Avcı is as heavy as Çakıl.  
D) Boncuk is the oldest of all.

12.

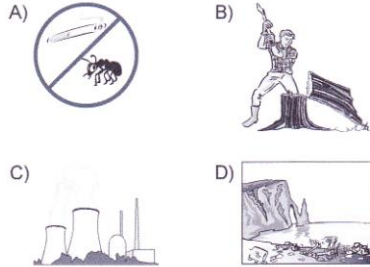


- A) He played the organ.  
B) He saw his cousin.  
C) He watered the grass.  
D) He injured his leg.

13.



Yukarıdaki uyarı panosuna ait olan resmin bulunduğu seçeneği işaretleyiniz.



14. Verilen resim seçeneklerden hangisinin reklamı olabilir?



- A) A cd player.  
B) An alarm clock.  
C) A new bicycle.  
D) A laptop computer.

15.

I am studying quietly in a large room.  
If you need a book, you can find here.  
Where am I ?



- A) In the library.  
B) In the corridor.  
C) In the canteen.  
D) In the supermarket.

## 8. SINIF

## YABANCI DİL TESTİ

A

BU GRUPTAKİ SORULAR YABANCI DİLLERE AİTTİR. OKUDUĞUNUZ YABANCI DİLE AİT SORULARI CEVAPLANDIRINIZ.

## İNGİLİZCE

1.- 9. sorularda boş bırakılan yerlere uygun dü-şen sözcük veya ifadeyi bulunuz.

1. Esen wants to study literature as .....

- A) she likes reading and writing.
- B) she is interested in science.
- C) she enjoys holidays.
- D) she is good at music.

2. Burak: It's important to be sympathetic to people.

Ceren: ..... I agree with you.

- A) I don't think so!
- B) That is not right!
- C) Absolutely!
- D) What a shame!

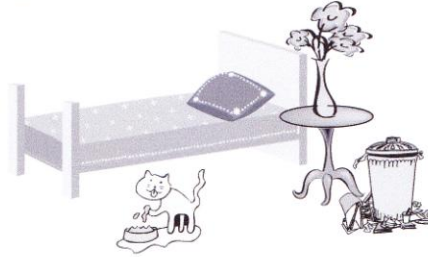
3. Gül : I have been very busy lately and I hate doing housework.

Mehmet : Don't worry! I'll help you.

Gül : Really? .....

- A) What a pity! Poor you!
- B) I'm sorry, I can't.
- C) That's an awful offer, isn't it?
- D) Ah, that's very kind of you!

4.



Serhat hasn't ..... yet.

- A) made his bed
- B) put the rubbish out
- C) fed the family pet
- D) bought some flowers

5.



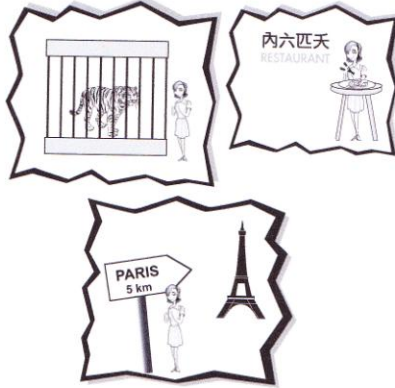
- A) sunburn
- B) sun rays
- C) sun cream
- D) sunglasses

8. SINIF

YABANCI DİL TESTİ

A

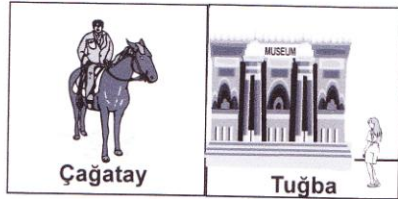
6. Melda is a thirteen year old girl. This is her photo album:



She's never .....

- A) driven a car.
- B) been abroad.
- C) eaten Chinese food.
- D) seen a tiger.

7.



Çağatay prefers ..... to visiting museums.

- A) mountain biking
- B) horse riding
- C) windsurfing
- D) parasailing

8.



- A) friendly
- B) tactful
- C) noisy
- D) punctual

9.



- A) You're joking!
- B) That's incredible!
- C) You look terrible!
- D) See you later!

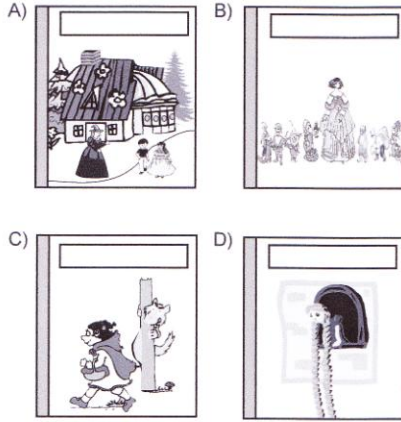
## 8. SINIF

## YABANCI DİL TESTİ

A

10. "Their father and stepmother leave Harry and Sally alone in the forest. There they see a little house. The owner of the house is an old witch ..."

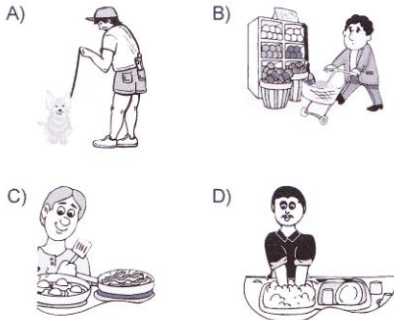
Yukarıda bir bölümü verilen hikâyenin kitap kapağı olabilecek seçeneği işaretleyiniz.



11.

	TO DO LIST
Ali	do the shopping
Bülent	make breakfast
Özkan	do the washing up
Burak	take the dog for a walk

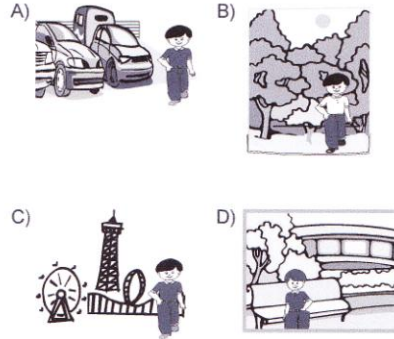
Yukarıda verilen tabloya bakarak Ali'nin bulunduğu resmi işaretleyiniz.



12. Konuşma balonunda verilen ifadeye uygun gelen resmi işaretleyiniz.

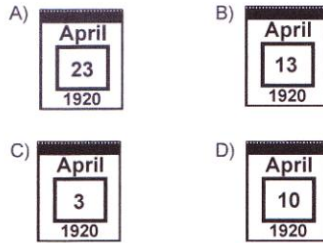
In my dream last night,  
I was playing at an  
amusement park.

?



13. M. Kemal Atatürk opened the Turkish Grand National Assembly on the twenty-third of April 1920.

Hangi takvim yaprağı yukarıda anlatılan olaya aittir?

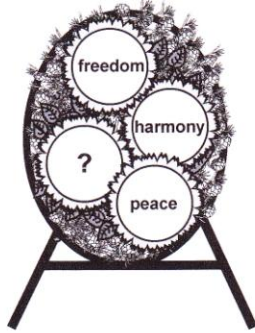


8. SINIF

YABANCI DİL TESTİ

A

14.



Barış çelenginin tamamlanması için soru işaretiyle gösterilen yere hangi çiçek getirilmelidir?

- A) B)
- C) D)

15. Someone stole a computer from Cumhuriyet Primary School yesterday. At the crime scene the police found some footprints. The thief was wearing trainers. They also found a black crayon. Who do you think stole the computer?

- A) B)
- C) D)

16. Sibel has a good sense of humour. She likes listening to jokes. Who can tell jokes for Sibel?

- A) B)
- C) D)

17.

Where should I put these empty milk bottles to recycle them?



- A) B)
- C) D)

**CURRICULUM VITAE**

**Name & Surname** : Tuğba AKINCI  
**Place of Birth** : Keçiborlu- ISPARTA  
**Date of Birth** : 1983  
**Graduation Degree** : Gazi University  
Education Faculty  
English Language Teaching Department  
**Date of Graduation** : 2005  
**Work Experience** : English Language Teacher  
(Since 2005)