

**People's Democratic Republic of Algeria**  
**Ministry of Higher Education and Scientific Research**  
**Mohamed Khider University of Biskra**  
**Faculty of Letters and Foreign Languages**  
**Department of Foreign Languages**  
**Section of English**



**INVESTIGATING LEARNERS' ATTITUDES TOWARDS  
THE IMPLEMENTATION OF A FORMAL ORAL TEST:  
The case of 4th year pupils at Cid Noureddine Middle School in Tolga.**

Dissertation submitted to the Department of Foreign Languages as a partial fulfilment  
of the requirements for the degree of Master in Language Sciences

**SUBMITTED BY:**  
**KHIDAR Khadoudj**

**SUPERVISED BY:**  
**Dr. HOADJLI Ahmed Chaouki**

**BOARD OF EXAMINERS:**

<b>Mr. CHENINI Abdelhak</b>	(Chairperson)	(University of Biskra)
<b>Dr. HOADJLI Ahmed Chaouki</b>	(Supervisor)	(University of Biskra)
<b>Dr. MEHIRI Ramdane</b>	(Examiner)	(University of Biskra)
<b>Miss. HADDAD Mimouna</b>	(Examiner)	(University of Biskra)

**Academic Year: 2016 - 2017**

## **Declaration**

I, **Khadoudj Khidar**, do hereby solemnly declare that the work I have presented in this dissertation is my own, and has not been submitted before to any other institution or university for a degree.

This work was carried out and completed at Mohamed KHEIDER University of BISKRA, ALGERIA.

Certified:

Miss: Khadoudj Khidar

Master Student, Section of English.

## **Dedication**

I dedicate this dissertation to:

My beloved mother Fatima, and my dear father Mohammed Salah

My sisters and my brother

My grandparents

My aunts and my uncles

My deepest gratitude for their support and encouragement.

I also dedicate this work to the soul of martyr Mohamed Kheider

## **Acknowledgements**

First and foremost, I would like to thank Allah the Almighty for giving me strength and patience to accomplish this modest work.

Special gratitude would go to my supervisor **Dr. HOADJLI Ahmed Chaouki** for his valuable guidance and assistance.

My thanks also go to the members of the jury **Dr. MEHIRI Ramdane, Mr. CHENINI Abdelhak**, and **Miss. HADDAD Mimouna** for their efforts to evaluate this work.

Finally, I would like to express my gratitude to my pupils who participated in this research study.

## Abstract

The ability to speak English effectively and appropriately is a priority for many English language learners. In Algeria, several reforms have been introduced to the English language curriculum to improve pupils' communication skills. However, there have been no changes in formal assessment. Achievement tests still assess vocabulary, grammar and some written sub-skills. Hence, it can be argued that they do not provide reliable and valid evidence of pupils' achievement and progress. Accordingly, the present research aimed at investigating pupils' attitudes towards the implementation of a Formal Oral Test in order to improve English language assessment and testing. In addition, it sought to determine how pupils perceived their experience of group speaking assessment. To address all these issues a qualitative research approach was adopted. Besides, the research strategy that was used was a case study. Concerning the data collection methods, we employed a preliminary questionnaire, a final questionnaire, and interview. The population of this study was fourth year pupils at Cid Noureddine Middle School in Tolga, which consisted of 111 pupils. Only 36 pupils were chosen to participate in this investigation. The findings revealed that many pupils had positive attitudes towards the implementation of a Formal Oral Test. This is mainly because pupils believe that their test scores do not reflect their level in the English language since tests do not assess their speaking abilities. In terms of their perceptions of the experience of the group format, the majority of the pupils said that it provided a comfortable context for speaking assessment. Moreover, we can say that the obtained results confirmed the hypotheses that this research was based on. Therefore, teachers are required to provide pupils with more opportunities to develop their communication skills. Also, policy makers in Algeria and test designers are recommended to consider the implementation of a Formal Oral Test to improve English language learning and assessment.

**Key Terms:** Achievement test, communication skills, formal assessment, formal oral test, group format, speaking abilities.

## List of Figures

**Figure 2.1** The Teaching-speaking cycle

p 35

## List of Tables

<b>Table 3.1:</b> The Order of using data collection methods	p 62
<b>Table 3.2:</b> Pupils' reasons for studying English language	p 64
<b>Table 3.3:</b> Pupils' awareness of the importance of all aspects of English language	p 64
<b>Table 3.4:</b> Pupils' perceptions of language skills	p 65
<b>Table 3.5:</b> Correspondence between test-contents and syllabus	p 65
<b>Table 3.6:</b> Testing methods variety	p 66
<b>Table 3.7:</b> The Tested ability	p 66
<b>Table 3.8:</b> Test-items distribution	p 67
<b>Table 3.9:</b> Neglected language skills in achievement tests	p 67
<b>Table 3.10:</b> Correspondence between test content and pupils' interest	p 68
<b>Table 3.11:</b> Correlation between test scores and pupils' level	p 68
<b>Table 3.12:</b> Pupils' writing abilities	p 69
<b>Table 3.13:</b> Pupils' speaking abilities	p 69
<b>Table 3.14:</b> Pupils' opinions about the importance of listening and speaking assessment	p 70
<b>Table 3.15:</b> Pupils' attitudes towards the implementation of a Formal Oral Test	p 71
<b>Table 3.16:</b> Pupils' favourite speaking assessment format	p 72
<b>Table 3.17:</b> Pupils' opinions about the English speaking test	p 72
<b>Table 3.18:</b> Pupils' experiences of the questions of the speaking test	p 73
<b>Table 3.19:</b> Pupils' opinions about questions and tasks of the speaking test	p 73
<b>Table 3.20:</b> Pupils' opinions of the assessor	p 74
<b>Table 3.21:</b> Pupils' attitudes towards group assessment	p 75
<b>Table 3.22:</b> Pupils' opinions about language skills	p 76
<b>Table 3.23:</b> Pupils' realisation of the importance of all aspects of English language	p 76
<b>Table 3.24:</b> Correlation between test scores and pupils' level	p 77
<b>Table 3.25:</b> Pupils' attitudes towards the implementation of a Formal Oral Test	p 78

## List of Abbreviations and Acronyms

**BEM** Examination: Brevet Examination

**CR:** Criterion-referenced

**DIF:** Differential Item Functioning

**Dr.:** Doctor

**DTF:** Differential Test Functioning

**EFL:** English as a Foreign Language

**et al.:** et alia - and others

**H:** Hypothesis

**Ibid:** Ibidem (Latin) - the same previous source -

**i.e.:** Id est - it means

**NR:** Norm-referenced

**p.:** Page

**Q:** Question

**S:** the standard deviation

**Vs:** versus

**X:** the mean

**&:** and

**%:** Percent



## **List of Appendices**

**Appendix 1:** The Preliminary questionnaire: English version

**Appendix 2:** The Preliminary questionnaire: Arabic version

**Appendix 3:** The Final questionnaire: English version

**Appendix 4:** The Final questionnaire: Arabic version

**Appendix 5:** The Interview

**Appendix 6:** The speaking test

## CONTENTS

<b>Declaration</b>	II
<b>Dedication</b>	III
<b>Acknowledgements</b>	IV
<b>Abstract</b>	V
<b>List of Figures</b>	VII
<b>List of Tables</b>	VIII
<b>List of Abbreviations and Acronyms</b>	IX
<b>List of Appendices</b>	X
<b>Contents</b>	
<b>General Introduction</b>	p 1
<b>1. Statement of the Problem</b>	p 1
<b>2. Research Questions</b>	p 2
<b>3. Research Hypotheses</b>	p 3
<b>4. Research Aims</b>	P 3
<b>5. Significance of the Study</b>	P 3
<b>6. Research Methodology</b>	P 4
<b>7. Literature Review</b>	P 5
<b>8. Structure of the Dissertation</b>	P 6
<b>CHAPTER ONE: LANGUAGE EVALUATION AND ASSESSMENT: REVISING THE RELATED REVIEW OF LITERATURE</b>	
Introduction	p 10
1.1 Evaluation, Measurement and Testing	P 10
1.1.1 Evaluation	P 10
1.1.2 Measurement	P 11
1.1.3 Testing	P 11
1.2 Types of Assessment	P 12
1.2.1 Formative assessment	P 12
1.2.2 Summative assessment	p 13
1.3 Purposes and Types of Tests	p 14
1.3.1 Test classification criteria	p 14
1.3.1.1 Intended use	p 15

1.3.1.2	Content	p 15
1.3.1.3	Frame of reference	p 15
1.3.1.3.1	Norm-referenced tests	p 16
1.3.1.3.2	Criterion-referenced tests	p 16
1.3.1.4	Scoring procedure	p 17
1.3.1.5	Testing method	p 17
1.4	Qualities of Language Test	p 19
1.4.1	Test Qualities	p 20
1.4.1.1	Reliability	p 20
1.4.1.2	Construct validity	p 20
1.4.1.3	Authenticity	p 20
1.4.1.4	Interactiveness	p 21
1.4.1.5	Impact	p 21
1.4.1.6	Practicality	p 21
1.5	Test Fairness	p 22
1.5.1	Validity	p 22
1.5.2	Absence of bias	p 22
1.5.3	Access	p 23
1.5.4	Administration	p 23
1.5.5	Social consequences	p 24
1.5.6	Implications for test development	p 24
1.6	Language Test Design and Development	p 25
	Conclusion	p 25
<b>CHAPTER TWO: THE SPEAKING SKILL: FUNDAMENTALS</b>		
	Introduction	p 28
2.1	Definition of Speaking	p 28
2.2	Components Underlying Speaking Effectiveness	p 30
2.2.1	Grammatical competence	p 30
2.2.2	Discourse competence	p 30
2.2.3	Sociolinguistic competence	p 30
2.2.4	Strategic competence	p 31
2.3	Teaching Speaking	p 31
2.3.1	Interaction as the key to improving EF L learners' speaking abilities	p 31
2.3.1.1	Small talk	p 31

2.3.1.2 Interactive activities	p 32
2.3.1.2.1 Aural: Oral activities	P 32
2.3.1.2.2 Visual: Oral activities.	p 32
2.3.1.2.3 Material-aided: Oral activities	p 33
2.3.1.2.4 Culture awareness: Oral activities	p 33
2.3.2 A Model for teaching speaking	p 34
2.3.2.1 Stage 1: Focus learners' attention on speaking	p 36
2.3.2.2 Stage 2: Provide input and / or guide planning	p 36
2.3.2.3 Stage 3: Focus learners' attention on speaking	p 37
2.3.2.4 Stage 4: Focus on language /skills /strategies	p 37
2.3.2.5 Stage 5: Repeat speaking tasks	p 37
2.3.2.6 Stage 6: Direct learners' reflection on learning	p 38
2.3.2.7 Stage 7: Facilitate feedback on learning	p 38
2.4 Assessing Speaking	P 39
2.4.1 The importance of rating in the speaking assessment cycle	P 40
2.4.2 Designing the rating process	p 40
2.4.3 Useful speaking scores	p 41
2.4.4 Speaking assessment procedures	p 41
2.4.4.1 Communication-oriented tasks and criteria	p 41
2.4.4.2 Pair and group tasks	P 42
2.4.4.3 Rating checklist	P 42
Conclusion	P 43
<b>CHAPTER THREE: FIELDWORK AND DATA ANALYSIS</b>	
Introduction	p 46
3.1 Research Methodology for this Dissertation: Theoretical Background	p 46
3.1.1 Research approach	p 46
3.1.1.1 Qualitative approach	p 47
3.1.1.2 Quantitative approach	p 47
3.1.1.3 A Mixed-methods approach	p 48
3.1.2 Research design / strategy	p 49
3.1.2.1 Qualitative designs	p 49
3.1.2.1.1 Narrative	p 50
3.1.2.1.2 Phenomenology	p 50
3.1.2.1.3 Ethnography	p 50

3.1.2.1.4 Case study	p 51
3.1.2.1.5 Grounded theory	p 51
3.1.2.2 Quantitative Designs	p 51
3.1.2.2.1 Surveys	p 51
3.1.2.2.2 Experiments	p 52
3.1.2.3 A mixed-methods design	p 52
3.1.2.3.1 Convergent parallel-mixed- methods design	p 52
3.1.2.3.2 Explanatory sequential-mixed- methods design	p 53
3.1.2.3.3 Exploratory sequential-mixed- methods design	p 53
3.1.3 Data collection methods	p 53
3.1.3.1 Qualitative data collection methods	p 54
3.1.3.1.1 Observation	p 54
3.1.3.1.2 Interviews	p 54
3.1.3.1.3 Questionnaires	p 55
3.1.3.1.4 Focus group	p 55
3.1.3.2 Quantitative data collection methods	p 55
3.1.3.2.1 Tests	p 55
3.1.3.2.2 Structured questionnaires	p 56
3.1.3.3 A Mixed-methods data collection methods	p 56
3.1.4 Sampling	p 57
3.2 Research Methodology for This Dissertation: Procedures	p 58
3.2.1 Research approach	p 58
3.2.2 Research design / strategy	p 58
3.2.3 Data collection methods	p 58
3.2.3.1 Questionnaires	p 59
3.2.3.1.1 Aim	p 59
3.2.3.1.2 Structure and content	p 59
3.2.3.1.3 Piloting and validation	p 60
3.2.3.2 Interview	p 61
3.2.3.2.1 Aim	p 61
3.2.3.2.2 Structure and content	p 61
3.2.4 Data collection procedures	p 61
3.2.4.1 Data collection proceduees for the qestionnaires	p 61
3.2.4.2 Data collection proceduees for the interview	p 62

3.2.5 Population and sampling technique	p 62
3.2.6 Data analysis procedures	p 63
3.2.6.1 Data analysis procedures for the questionnaires	p 63
3.2.6.2 Data analysis procedures for the interview	p 63
3.3 Results	p 64
3.3.1 Pre-questionnaire	p 64
3.3.2 Post-questionnaire	p 72
3.3.3 Interview	p 78
3.4 Discussion of the Results	p 81
3.5 Synthesis of the Findings	p 84
Conclusion	p 85
<b>General conclusion</b>	p 86
<b>References</b>	p 88
<b>List of Appendices</b>	
<b>Glossary</b>	
<b>ملخص</b>	

## **General Introduction**

Over the last decades, there has been an increasing interest in learning English language all over the world. The main reason behind this is that English language has become the main means of communication in all domains of life. Therefore, many educational programmes have been devised to help learners master this foreign language. However, since the main purpose of learning a foreign language is communication, more attention has been given to teaching and assessing the speaking skill.

The ability to speak English language effectively and appropriately is now a priority for many foreign language learners. Learners very often evaluate their success in English language learning on the basis of how well they feel they have improved in their spoken language proficiency. Hence, speaking is considered as the most important language skill that requires more consideration.

In Algeria, several changes have been introduced to the curriculum to improve the level of pupils in the English language, and to enhance their speaking abilities. These reforms aim at providing pupils with more opportunities to speak and practise English language. Nevertheless, the adoption of these changes is hardly noticed. Teachers still focus on teaching vocabulary and grammatical rules. We believe that the root cause of this is that there have been no crucial changes in the assessment system. In other words, the way teachers assess their pupils has not changed yet.

Consequently, this research will be conducted to investigate pupils' attitudes towards the implementation of a Formal Oral Test to enhance English language assessment and testing. Besides, we will attempt to explore their views on group speaking assessment.

### **1. Statement of the Problem**

English language has increasingly become the main means of communication around the world. As a result, teaching English language as a foreign language has captured attention and interest. Despite the fact that learning a foreign language requires the mastery of the four skills, the speaking skills were the least considered and practised. In recent years, however, more attention has been given to the improvement of EFL learners' speaking abilities since the main purpose of learning a foreign language is communication. Thus, teaching and effectively assessing EFL learners' speaking skills have become a priority in many countries.

In Algeria, several changes have been introduced to the English language curriculum to enhance pupils' speaking abilities. Although the focus was on teaching grammar and vocabulary, teachers are now urged to take into account all aspects of English language learning, as well as to develop pupils' speaking skills. Also, there has been a shift towards the use of different interaction patterns, such as individual, pair and group work, to provide pupils with enough opportunities to practise English. Furthermore, for the last couple of years, a tutorial class has been brought to the curriculum, and teachers have been instructed to devote it to group tasks and oral communication activities.

However, there have been no considerable changes in formal assessment. The currently used tests still assess grammar, vocabulary and some written sub-skills. They do not assess pupils' speaking abilities. Thus, they do not provide reliable and valid scores. In other words, they do not present reliable and valid evidence of pupils' achievement and progress. Besides, the way which is used to test pupils has led to the slow adoption of the new policy and techniques which focus on the development of pupils' communication skills. Moreover, it has affected the way pupils perceive English language learning and assessment.

For all these reasons, we believe that the implementation of a Formal Oral Test is now a necessity to enhance English language assessment and testing. That is, more attention must be given to speaking assessment to obtain useful test scores that reflect pupils' level in the English language. This will also motivate teachers to change their practices in the classroom, and create more opportunities for pupils to practise English language and improve their speaking abilities.

## **2. Research Questions**

This study investigates the following research questions:

- Q1:** What are pupils' perceptions of English language learning and assessment?
- Q2:** What are pupils' attitudes towards the implementation of a Formal Oral test; and what are the factors that contribute to their attitudes?
- Q3:** How do pupils perceive their experience of group speaking assessment?



### **3. Research Hypotheses**

This research is based on the following hypotheses:

**H1:** There is a relationship between pupils' perceptions of English language learning and assessment and their attitudes towards the implementation of a Formal Oral Test.

**H2:** There is a positive relationship between the use of group speaking assessment and pupils' attitudes towards the implementation of a formal oral test.

**H3:** There is a positive relationship between the use of group speaking assessment and pupils' speaking performances.

### **4. Research Aims**

1-The general aim of the present study is to investigate pupils' attitudes towards the implementation of a Formal Oral Test, as well as to identify the factors that contribute to their attitudes.

2- The specific aims are:

- to explore pupils' perceptions of English language learning and assessment;
- to determine the effectiveness of group speaking assessment; and
- to see how pupils perceive their experience of group format.

### **5. Significance of the Study**

The present research mainly aims at contributing in the improvement of English language assessment and testing. It seeks to sensitise policy makers and test developer to the importance of speaking tests. It is so important to discover whether or not the implementation of a Formal Oral Test can enhance English language assessment and testing, and create a balanced assessing system which covers all language aspects and the four language skills.

Similarly, the information that will be gathered in this study will be useful to test developers as this research aims at providing evidence of the effectiveness of group speaking assessment. The findings of this study can also provide insight about how pupils perceive English language learning and assessment.

Moreover, this research can be beneficial to educators and teachers because it sheds light on useful ways to facilitate oral communication and assessment. Also, it can draw their attention to the fact that speaking is the most important skill, and that it becomes a

priority for pupils to develop their speaking abilities. Hence, teachers should give more attention to teaching the speaking skills.

Finally, this study can be useful to researchers in the language testing field as it can provide them with useful information to conduct further studies in order to enhance English language assessment and testing.

To sum up, the present study on assessing the speaking skills can change the perceptions and practices of English language learning and assessment.

## **6. Research Methodology**

As it is mentioned before, the main purpose of this research study is to investigate pupils' attitudes towards the implementation of a Formal Oral Test, and to identify the factors that contribute to their attitudes. Besides, it aims at determining the effectiveness of group speaking assessment. Finally, it seeks to explore pupils' perceptions of English language learning and assessment.

To better tackle these issues, the research approach that will be used in this study will be a qualitative research approach because of its suitability to answer the research questions as it offers the appropriate data collection methods. In addition, the research strategy that will be adopted is a case study. With regard to the data collection methods, we will use a questionnaire and interview to collect the appropriate data. A preliminary questionnaire will be used to explore pupils' perceptions of English language learning and assessment, whereas a final questionnaire will be employed after a speaking test to investigate pupils' attitudes towards the implementation of a Formal Oral Test. We also intend to conduct a semi-structured interview to elicit more information about the participants' attitudes and opinions.

The target population of this study is fourth year pupils at Cid Nouredine Middle School in Tolga, Biskra (Algeria). It consists of 111 pupils. One class, which includes 36 pupils, has been chosen to participate in this investigation. Therefore, the sampling technique which has been selected for the present research is a purposive sample. It is important to note that the participants will be divided into groups of four during the speaking test which they will take in this study. Then, we will question them about their attitudes towards the implementation of a Formal Oral Test, as well as their opinions of group speaking assessment.

## **7. Literature Review**

Although speaking a foreign language is the most valuable language skill, testing this skill was not taken into consideration and was almost neglected (Flucher, 2014). However, since English language has become the lingua franca in our world today, oral communication has recently gained an increasing attention. More importance has been given to teaching and effectively assessing EFL learners' communicative speaking skills (Kwon, 2015).

Therefore, there has been a dissatisfaction with the quality of information that can be collected from the traditional tools of assessment. It has been argued that written tests too often focus on recognition or recall of disconnected skills and concepts. Hence, many educators suggest that learners have to carry out authentic assessment tasks or 'performance tasks' (Ryan & al., 2013). With regard to Algeria, many teachers believe that the currently used tests have negative effects on the content of their teaching and classroom assessment. They argue that these written tests focus more on grammar, vocabulary and some written sub-skills without giving attention to the speaking skills. Consequently, it is proposed to adopt an incorporated oral test in the achievement test (Hoadjli, 2016).

In recent years, researchers, in the language testing field, have been investigating the effects that an oral test can have on the teaching learning process, and whether they can devise a valid and reliable oral test. It has been found that an oral test can have interesting and positive effects on teaching strategies and materials (ibid). Also, other studies were carried out to explore the best ways to assess the speaking skills. For example, Kwon (2015), who conducted a study on paired speaking assessment, claims that students perceived paired format as effective and providing a comfortable context for speaking assessment. In addition, he investigated students' attitudes towards Formal Oral Test. The findings showed that learners varied greatly in their attitudes. Some agreed stating that it would be a considerable step towards improving the educational system. Others disagreed and cited increased pressure and anxiety, decreased confidence, and competitive drive as potentially negatively affecting their speaking performance.

Learners can get a lot of benefits from speaking assessment. For example, learners, who have difficulties in communicating their ideas on paper, may find communicating their ideas orally is much easier to do. The reason behind this is that they do not have to take into consideration the punctuation marks. Despite the fact that writing is an important and essential skill, allowing learners to convey their ideas during group activities is a good way

to help them 'enjoy' (Casas, 2011).

In conclusion, the implementation of a Formal Oral Test becomes an increasing necessity. It is a crucial step to enhance English language assessment since written tests are no more considered as reliable and valid assessment tools. However, we think that there are other questions which can be raised. For example, 'when should teachers start testing EFL learner' speaking abilities?' That is, at what age EFL learners should start taking a Formal Oral Test. Also, there are other questions which are related to speaking assessment procedures. For instance, 'Should we test learners individually, in pairs or in groups?' i.e., 'What is the best way to assess EFL learners' speaking skills?'

## **8. Structure of the Dissertation**

This dissertation consists mainly of two main parts which are the theoretical and practical parts. It is divided into three chapters. The first and second chapters are related to a related literature review; whereas, the third one is devoted to the fieldwork and data analysis.

### **Chapter One**

This chapter displays a theoretical overview of language evaluation and assessment. Initially, it provides definitions of three basic terms, namely assessment, evaluation and testing. Then, it identifies the purposes and types of tests, as well as the criteria used to classify them. In addition, it presents the qualities of language test, and discusses the concept of test fairness. Finally, the chapter tackles language test design and development.

### **Chapter Two**

This chapter mainly deals with teaching speaking and assessment. First, it states the components underlying speaking effectiveness and reviews the communicative competence proposed by Canale and Swain. Additionally, it displays a model for teaching speaking which can help improve learners' speaking competence. Furthermore, it tackles the difficulties that teachers face when teaching this basic skill, and how interaction plays an important role in enhancing EFL learners' speaking abilities. Finally, the chapter discusses speaking assessments and the procedures used to rate performances.

### **Chapter Three**

This chapter is devoted to fieldwork and data analysis. First, it reviews the relevant research methodology in our study. In addition, it displays the procedures used to conduct the present study. Moreover, it presents the results obtained using two data collection methods, namely a questionnaire and interview. Finally, it discusses and synthesise the findings, and draw some conclusions related to English language learning and assessment.

**CHAPTER ONE :**

**LANGUAGE EVALUATION AND ASSESSMENT:  
REVISITING THE RELATED REVIEW OF**

## **CHAPTER ONE: LANGUAGE EVALUATION AND ASSESSMENT: REVISING THE RELATED REVIEW OF LITERATURE**

Introduction	p 10
1.1 Evaluation, Measurement and Testing	P 10
1.1.1 Evaluation	P 10
1.1.2 Measurement	P 11
1.1.3 Testing	P 11
1.2 Types of Assessment	P 12
1.2.1 Formative assessment	P 12
1.2.2 Summative assessment	p 13
1.3 Purposes and Types of Tests	p 14
1.3.1 Test classification criteria	p 14
1.3.1.1 Intended use	p 15
1.3.1.2 Content	p 15
1.3.1.3 Frame of reference	p 15
1.3.1.3.1 Norm-referenced tests	p 16
1.3.1.3.2 Criterion-referenced tests	p 16
1.3.1.4 Scoring procedure	p 17
1.3.1.5 Testing method	p 17
1.4 Qualities of Language Test	p 19
1.4.1 Test Qualities	p 20
1.4.1.1 Reliability	p 20
1.4.1.2 Construct validity	p 20
1.4.1.3 Authenticity	p 20
1.4.1.4 Interactiveness	p 21
1.4.1.5 Impact	p 21
1.4.1.6 Practicality	p 21
1.5 Test Fairness	p 22
1.5.1 Validity	p 22
1.5.2 Absence of bias	p 22
1.5.3 Access	p 23
1.5.4 Administration	p 23
1.5.5 Social consequences	p 24

1.5.6 Implications for test development	p 24
1.6 Language Test Design and Development	p 25
Conclusion	p 25



# **CHAPTER ONE: LANGUAGE EVALUATION AND ASSESSMENT: REVISITING THE RELATED REVIEW OF LITERATURE**

## **Introduction**

This chapter is devoted to language evaluation and assessment. First, it reviews the differences between three basic terms which are assessment, evaluation and testing. It also provides definitions of the two major types of assessment, namely formative and summative assessment. Then, it identifies the purposes and types of tests, the criteria used to classify them, and their definitions. In addition, this chapter displays the qualities of language test, and discusses the concept of test fairness and the subsequent qualities which constitute test fairness framework. At the end, the chapter tackles language test design and development, and explains the primary and principal step in the development of language tests.

### **1.1 Evaluation, Measurement and Testing**

Language assessment is an integral part of foreign language teaching and learning process. It plays a significant role in the education of English language learners. Gitsaki and Coombe (2016) emphasise that whatever the type of language assessment is, it always involves collecting, interpreting and evaluating information about students' learning. As there is no single technique that can be employed to gather such kind of information, different assessment and evaluation tools are used to identify students' needs and enhance instruction.

According to Bachman (1990), the terms evaluation, measurement and test are often used interchangeably. However, this may distract attention from the fact that each concept has its distinctive features and that a clear understanding of the basic differences among them is crucial to 'the proper development and use of language tests.' (p. 18)

#### **1.1.1 Evaluation**

Evaluation can be defined as the judgements teachers make, as well as the actions they take which are based on the systematic information they collect about their learners, such as decisions related to students' promotions or retention at the end of the school year. It

also refers to the decisions teachers make regularly about the type of test to create, when to give it, and what language material or abilities to assess (Casas, 2011). In another definition, Bachman (1990) considers evaluation as the gathering of pertinent and reliable information, such as verbal descriptions and comprehensive impressions in order to evaluate learners.

### **1.1.2 Measurement**

Casas (2011) defines measurement as 'all the information that teachers collect as part of the assessment process' (p. 182), such as responses in classroom, test scores, homework and performance assessment. However, Orlich et al. (2013) emphasise that measurement is a quantitative description of students' performance. In other words, it is the process of assigning numbers to assessment results. According to Bachman (1990), measurement can be defined as 'the process of quantifying the characteristics of persons according to explicit procedures and rules' (p. 18). Thus, measurement has three basic features which are: First, quantification which refers to the process of gathering quantitative or numerical data rather than qualitative descriptions, such as labels, for example, excellent, very good, good and average. However, non-numerical data can be converted to numbers in order to be analysed and interpreted and, as a result, they can form measurement. Second, characteristics of persons that can be divided into two kinds which are physical traits, such as weight and height, and most importantly mental abilities, such as intelligence, aptitude, achievement in reading comprehension and fluency in speaking. Third, explicit rules and procedures which indicate that the assignment of numbers must not be done randomly because it cannot be considered as measurement. Therefore, test, rating scales and ranking are among the various kinds of measures that can be used.

### **1.1.3 Testing**

'Testing is a universal feature of social life' (McNamara, 2000, p. 3). That is, tests have always been used to assess and provide evidence of people's abilities and capacities. Bachman (1990) illustrates that a test is one type of measurement tools which is devised to obtain a representative sample of an individual's behaviour. It is a way to assign numbers to the attributes and abilities of learners using an explicit procedure. In other words, tests are created to measure students' learning and achievement on the basis of specific

objectives and material which are already determined by the teacher (Ryan et al., 2013).

It can also be defined as a set of questions which students must answer within a limited period of time and under the same circumstances. Thus, it is considered an indicator of students' learning and a source of data which is used to rank students (cited in Orlich et al., 2013).

Green (2014) defines a test as an event that is arranged to obtain information about test takers' performance in order to make judgments about their knowledge, skills or abilities. It usually occurs in a fixed period of time, that is decided in advance, during which test takers or testees have to respond to a prompt, such as oral questions or a set of choices on a test paper. Their responses are measured using a mark scheme or a predetermined procedure which results in scores or grades that are officially recorded.

Therefore, it can be deduced that tests can have different forms and purposes and they can be oral or written, but in all cases they provide numerical data about students' abilities and attributes.

## **1.2 Types of assessment**

There are mainly two categories of assessment which are formative and summative assessment. The former is used to guide students' future learning whereas the latter is used to assess their achievement. Actually, both types serve both purposes. However, Newton (2007) argues that although formative and summative assessment may have the same form, the difference is mainly in the ways the results are interpreted and used (cited in Long et al., 2011).

### **1.2.1 Formative assessment**

Formative assessment is defined as the type of 'assessment that is used to monitor learning progress during instruction' (Ryan et al., 2013, p. 204). That is, it is the kind of assessment which occurs during the teaching learning process and which is used to evaluate students' learning as they progress through their education. It is used to determine what a student does or does not understand. Benjamin (2013) acknowledges that formative assessment concerns both teachers and students. It helps teachers discover their students' strengths and weaknesses, so that they can address students' needs and plan appropriate learning experiences to improve students' knowledge, skills and abilities. Therefore, it

serves as a means to enhance students' learning, as well as teachers' teaching. Two main points have to be considered so that formative assessment can be effective. First and foremost, teachers must know exactly what formative assessment process is, and its relationship to instructional scaffolding. They have to provide students with explicit feedback to improve their learning. They also have to adapt their way of teaching to meet students' needs. Second, teachers have to be good at using the different assessment techniques, such as observation, portfolios, performance tasks, students self-assessment and feedback (Darling-Hammond & Bransford, 2005).

### **1.2.2 Summative assessment**

Summative assessment is the type of assessment that is developed to measure students' achievement using grades. It usually occurs at the end of an instructional unit of study (Ryan et al., 2013). Consequently, it has an 'air of finality' (Benjamin, 2013, p. 3). Another definition is that of Long et al. (2011) which states that summative assessment includes formal tests and examinations which are often carried out at the end of a block of teaching. They involve assessing students' knowledge, skills and abilities in accordance with a specific curriculum. Therefore, they are so important to students, as they provide them with access to higher levels of education or to employment. Also, they are used to evaluate schools and teachers' performance; consequently, they are also known as 'high-stakes assessment'.

Therefore, this can lead to what is called curriculum backwash, whereby the content of tests govern what is taught or classroom practices. Long et al. (ibid) claim that this is not necessarily a bad thing since a test cannot provide a realistic assessment of performance across the whole curriculum. However, we partly disagree with this viewpoint because if we consider the status of English language teaching in Algeria, it can be noticed that formal tests and examinations focus mainly on certain aspects of English language, namely grammar and vocabulary. They give a little importance to reading and writing skills; however, they almost neglect the speaking skills. Consequently, we believe that formal test should cover the four skills and all aspects of language and that the various parts of test should be similar in content.

Moreover, research findings in fields of cognition, motivation and measurement suggest that students should demonstrate their proficiency in various ways and that they should be assessed on the basis of performance expectations instead of being assessed against each

other (Darling-Hammond & Bransford, 2005).

### **1.3 Purposes and Types of Tests**

Allison (1999) believes that any test has three main parameters which are content, method and purpose. According to this author, classroom language testing, whether formal or informal, mainly serves as a means to assess students' learning, and to determine what has been learnt and what requires remedial instruction. Besides, it is used for many other purposes. He demonstrates that the labels of some kinds of language test can indicate their purposes while others do not. However, recognising the difference between the types of tests and their purposes is not quite easy, as there are certain types whose names do not provide any hints of their purposes, such as multiple-choice test.

A test label may convey the general purpose of the test, the content, an aspect of method, or a combination of these three parameters. However, its main purpose depends on assessment procedures which are developed according to the needs of testers, test takers and test users. It is also important to emphasise that the selection of test method is not completely unrelated to the principal purpose of a test (ibid).

According to Henning (1987), the purposes of tests can be diagnosis and feedback, screening and selection, placement, programme evaluation, providing research criteria, or assessment of attitudes and sociopsychological differences (cited in Allison, ibid). Concerning the types of language tests, there are many kinds. These are some of them:

- Objective Vs. subjective tests;
- Direct Vs. indirect tests;
- Discret-point Vs integrative tests;
- Aptitude, achievement, and proficiency tests;
- Criterion- or domain-referenced Vs. norm-referenced or standardised tests; and
- Speed tests Vs. power tests.

#### **1.3.1 Test classification criteria**

Bachman (1990) argues that language tests can be classified according to five various features which are: the purpose or the intended use, the content, the frame of reference, the scoring procedure, and the testing method.

### **1.3.1.1 Intended use**

Bachman (ibid) states that the purpose is the principal aspect to consider either when designing a language test or when interpreting its results. In other words, any language test is created according to a predetermined use. Hence, since language tests are used to gather systematic information in order to make a wide range of decisions, they can be classified on the basis of those decisions as follows:

- In terms of admission decisions, there are selection, entrance and readiness tests.
- With regard to identifying appropriate instructional level and remedial instruction, there are placement and diagnostic tests.
- In terms of learners' educational progress, there are progress, achievement, attainment or mastery tests.

### **1.3.1.2 Content**

Bachman (ibid) illustrates that the content of language tests can be developed in accordance with a specific domain of content or a theory of language. If language tests are based on syllabuses, they are referred to as achievement tests. If they are based on a theory of language, they can be divided into two types. First, language aptitude tests which are based on a language acquisition theory. Second, proficiency tests which are based on a language use theory.

### **1.3.1.3 Frame of reference**

The interpretation of the results of language tests can be carried out in two various ways in accordance with the adopted frame of reference. The first way is referred to as norm-referenced interpretation, whereby test scores are interpreted on the basis of a given group performance. The second way is referred to as criterion- or domain-referenced interpretation, whereby test results are interpreted in relation to a specific level or domain of ability. Thus, there are two types of tests which are norm-referenced and criterion-referenced tests (ibid).

### **1.3.1.3.1 Norm-referenced tests**

As mentioned before, the test scores of norm-referenced tests are interpreted with respect to the performance of a particular norm group, which is a large group of individuals that is regarded as a representative of the target population. As a result, they result in 'normative' interpretations. A NR test is given to the norm group whose performance characteristics, which are generally the mean  $X$  and the standard deviation  $S$ , are then used as reference points for interpreting other test takers' performance. However, NR results are sometimes interpreted and reported with respect to the target group, not to a separate norm group.

To obtain results that can be most easily interpreted, NR tests, which are sometimes called psychometric tests, are developed to maximise the differences among individuals in a particular group.

The standardised test is regarded as the typical NR test. It has three features. First, standardised tests have a fixed content which may be related to either a language proficient theory or to a description of language users' expected needs. Also, they have uniform procedures for administration and scoring. Finally, they have been scientifically examined and tested and their characteristics are well-known (ibid).

### **1.3.1.3.2 Criterion-referenced tests**

Criterion-referenced tests allow test users to interpret a test score according to a criterion level of ability or domain of content. When developing a CR test, there are two main points to consider. First, a CR test has to sufficiently reflect the criterion ability level or the content domain. Also, it has to be sensitive to levels of ability or degrees of mastery of the various constituents of that domain.

In conclusion, the main differences between NR and CR tests are in their design, construction, and development, as well as in the scales they produce and their interpretation. However, it is important to point out that these frames of reference are not reciprocally exclusive. For example, a NR test can be scaled to an existing CR test (ibid).

#### **1.3.1.4 Scoring procedure**

Almost all aspects of tests require making subjective decision, such as identifying the content to be covered or deciding on the best ways to to construct the test items. Even test takers have to make subjective decisions related to the best ways to answer the questions. However, as Pilliner (1968) states, what differs objective tests from subjective tests is the scoring procedure (cited in Bachman, *ibid*). With regard to objective tests, a test taker's answer is considered to be correct with reference to predetermined criteria, which is, for example, the case for the mutiple-choice technique. However, with reference to subjective tests, the response is judged according to the scorer's subjective interpretation of the scoring criteria, which is, for example, the case for oral interview.

#### **1.3.1.5 Testing method**

As there is a broad range of different testing methods that have been devised and will, probably, be devised, it is not possible to make a limited list of the methods used for language tests. However, there is one wide type which has been considered by language testers that is 'performance test' which is used to predict test takers' performance in real life situations. Yet, performance tests, such as oral interview and essay, and the more frequently used methods, for instance, the multiple-choice, completion, dictation and cloze, are not regarded as single methods. They comprise different facets, such as instruction, types of input and task types, that are considered as a more precise way of classifying tests (*ibid*).

To sum up, Bachman (*ibid*) points out that the five features, discussed before, serve as a means to classify any given test and that the description of language tests from only one angle are likely to be misunderstood. He adds that although the parameters are distinct, they overlap in some areas. First, the terms 'achievement', 'attainment' and 'mastery' may refer to the type of content, as well as to the type of decision to be made. Also, the scores of proficiency and aptitude tests have been interpreted according to some reference group in spite the fact that there is no link between test content and frame of reference. Moreover, achievement tests are generally criterion-referenced. However, a knowledge of the different features used to describe tests can help understand how tests are distinguished from each other and provide tests users with ways to specify the characteristics of the tests they use. Finally, these features identify the approaches used to examine and demonstrate



reliability and validity.

According to McNamara (2000), language tests can be mainly distinguished on the basis of their methods and purposes; that is, 'how they are designed' and 'what they are for'.

With regard to testing method, tests can be basically divided into two types which are paper-and-pencil language tests and performance tests.

Paper-and-pencil language tests refer to the traditional and common form of test which requires a question paper. They serve mainly two purposes which are measuring one aspect of language, such as grammar or vocabulary, or assessing listening or reading comprehension. However, they are not frequently used to assess speaking and writing skills as they are generally in a fixed response format, such as multiple-choice format.

Performance based tests refer to language skills assessment in communicative tasks. They are, unlike the first type, very often used to assess the productive skills by eliciting a sample which is related to real life situations and which is assessed using a rating procedure.

On the other hand, language tests also vary according to their purposes. The purpose can have an impact on the form of test. Also, the same form can serve a range of various purposes. In this case, the main distinction is made between achievement tests and proficiency tests.

Achievement tests are related to the teaching learning process. They are used to collect data during or at the end of a language course to determine what has been and what has not been learnt. Consequently, they can have many forms, such as observational procedures which are based on classroom participation and activities and which are used to monitor progress, portfolio assessments, or end of course tests.

The use of multiple-choice format in developing achievement tests has been strongly criticised. It is believed that multiple-choice standardised tests have a negative effect on classroom practices in a way that they make teachers teach only what students will be tested on. However, during the last decades, the focus has been put on language use or language performance; as a result, the test does not correspond to the curriculum.

In fact, this is the case in Algeria where teachers still give great importance to grammar and vocabulary in their classrooms, and neglect the speaking skill because it is not assessed in achievement tests.

Therefore, an achievement test can be characterised as being 'self-enclosed' because it mainly assesses grammar and vocabulary knowledge and has no relation with language use in the external world. According to McNamara (ibid), this problem can be solved only if

the syllabus itself emphasises language use. However, we believe that this is not enough and there should also be crucial changes in the test format. This, in turn, will have an impact on the adopted teaching methodology.

McNamara (ibid) sees that achievement tests can be easily innovated according to the curriculum. They are influenced by the improvements in language assessment which are known as alternative assessment. With regard to this new approach, there must be a relationship between assessment, the aims of the curriculum, and instruction. Also, teachers may promote students' self-assessment; in other words, they may encourage students to assess and judge their own abilities in performance in different contexts.

Unlike achievement tests which assess students' previous knowledge, proficiency tests 'look to the future situation of language use' (ibid) which is referred to as the criterion. In other words, they predict students' performance in real life situations without considering any previous formal instruction. Recently, the characteristics of the criterion or language performance have been progressively taken into account when developing tests.

#### **1.4 Qualities of Language Test**

According to Bachman and Palmer (1996), the major characteristic of a test is usefulness. They state that a language test usefulness forms a kind of 'metric' which can evaluate all aspects of a test development and use. Hence, these two authors suggest a model which comprises six main qualities, along with three principles that are believed to form all together a foundation to judge the usefulness of a test. The three principles are:

- Principle 1: It is the overall usefulness of the test that is to be maximized, rather than the individual qualities that affect usefulness.
- Principle 2: The individual test qualities cannot be evaluated independently, but must be evaluated in terms of their combined effects on the overall usefulness of the test.
- Principle 3: Test usefulness and the appropriate balance among the different qualities cannot be prescribed in general, but must be determined for each specific testing situation.

(Bachman & Palmer, 1996, p. 18)

Bachman and Palmer (ibid) explain that for any language test to be useful, it must be developed with reference to a specific purpose, a particular group of test takers and a target language use. Therefore, test usefulness cannot be judged in a general way. It can be described in terms of six test qualities. However, specifying the appropriate or acceptable balance that should be among the different qualities is not possible. Evaluating the overall

usefulness of a particular language test is mainly subjective. For example, a test developer may want to design a test in a way that provides high degrees of reliability and validity, or a teacher may want to use a test that guarantees high levels of authenticity, interactiveness and impact.

## **1.4.1 Test Qualities**

### **1.4.1.1 Reliability**

It is often defined as consistency in measurement. That is, the same group of students should get the same, or nearly the same, grades if they take the same test but in different situations. They should also get the same grades whatever the version of a test they take. Similarly, students should obtain the same grades whoever the corrector is. In other words, the person who marks a test should make no difference and the scores have to be consistent.

### **1.4.1.2 Construct validity**

This quality has to do with the appropriateness and meaningfulness of the interpretations of a test scores. It refers to the extent to which a given test score can be considered as an indicator of the abilities which have to be tested with reference to a specific domain of generalization. According to Saraswati (2004), to characterise a test as being valid it has to measure what is supposed to measure. In fact, there are various types of validity, such as face validity and content validity. The former implies that a test must be neither too easy nor too difficult. It also does not to be unrealistic. That is, it has to appear in a reasonable way to assess learners. The latter means that a test has to be a 'balanced sample'; in other words, it has to appropriately assess all the areas it is supposed to test.

### **1.4.1.3 Authenticity**

It refers to the extent to which the features of a certain language test task match the characteristics of a target language use task. It is an important test quality mainly because of two reasons. First, it provides a link between test performance and the TLU domain. Second, it has a significant impact on the way test takers perceive a test which, in turn, can positively affect their performance (Bachman & Palmer, 1996).

#### **1.4.1.4 Interactiveness**

It is defined as the extent and the ways in which a test taker's language ability, topical knowledge and affective schemata are involved in performing a test task. As a result, it exists in the interaction between the language user or test taker and the task (ibid).

#### **1.4.1.5 Impact**

It is the impact that tests have on educational systems and society as well as on the individuals which constitute those organisations. Thus, the impact of test use can be on a micro-level, when referring to the individuals, as well as on a macro-level, when referring to the educational system or society as a whole (ibid).

#### **1.4.1.6 Practicality**

This quality differs in nature from the previous five qualities. Unlike the other qualities which are related to the uses of test scores, practicality is mainly concerned with the implementation of the test and whether it will be developed and used in the first place. In other words, the test will be impractical if the resources available are not enough to implement it. Thus, practicality refers to 'the relationship between the resources that are required in the design, development, and use of the test and the resources that will be available for these activities'. (ibid, p. 36)

Practicality is as important as the other qualities of usefulness because it affects the decisions to be made at every level. Test developers have to create a perfect balance among the qualities of reliability, construct valid, authenticity, interactiveness and impact. Moreover, the resources required to create this balance must be established with reference to the resources which are available. Therefore, the practicality of a particular test can be determined on the basis of two factors. First, the resources required to develop an operational test that can achieve the desired balance among the qualities. Second, the distribution and the management of the resources that are available, which are mainly human resources, material resources and time.

To sum up, these six qualities are complementary and test developers have to create a suitable balance among them.

## **1.5 Test Fairness**

Test fairness is a concept which is related to test evaluation. In fact, there is no coherent framework that can be used for evaluating tests and testing practices. However, Kunnan (2000) presents a Test Fairness framework that includes five basic test qualities which are validity, absence of bias, access, administration, and social consequences. Therefore, this framework considers fairness as a whole system of testing practice.

### **1.5.1 Validity**

According to Kunnan (ibid), there are three types of validity of a test score interpretation. First, content representativeness or coverage which refers to 'the adequacy with which the test items, tasks, topics and language dialect represent the test domain' (Milanovic & Weir, 2001, p. 37). This type is also called content validity. Second, construct or theory-based validity which can be defined as the quality when the test items, tasks, topics and language dialect sufficiently represent the construct, theory or underlying trait to be tested. Third, criterion-related validity which is related to the question of whether the test scores satisfy criterion variables, such as school or college grades. Finally, reliability which can be defined as the consistency of test scores. It can be divided into four types as follows:

- consistency of scores in different testing situations which is known as stability evidence;
- between two or more various versions of a test which is known as alternate form evidence;
- between two or more raters which is known as inter-rater evidence; and
- in the way test items measure a construct functions which is known as internal consistency evidence (ibid).

### **1.5.2 Absence of bias**

Absence of bias in a test means the absence of offensive language, unfair penalisation based on test taker's background, and disparate impact and standard setting. Firstly, offensive content or language refers to the content that is offensive to test takers from different backgrounds. For example, stereotype and insults, whether straightforward or covert, which are based on gender, age, race and ethnicity, religion, native language and national origin. Secondly, unfair penalisation based on test taker's background is the content that may lead to unjust punishment due to a test taker's gender, age, race and

ethnicity, religion, native language and national origin.

Finally, disparate impact and standard setting is the type of bias that refers to differing performances and resulting outcomes by test takers of different genders, races, religions, ages, native languages, and national origins. Such differences, which are known as salient test-taker characteristics, on test tasks and sub-tests should be examined for Differential Item/Test Functioning (DIF/DTF). Also, to ensure that a test does not predict success for one group than for another, a differential validity analysis should also be carried out. As regard to standard-setting, test scores should be analysed on the basis of the criterion measure and selection decisions. Hence, these analyses can ensure to test developers and test users that group differences are linked to abilities that are being tested rather than to other unrelated factors (ibid).

### **1.5.3 Access**

With regard to his quality, Kunnan (2000) differentiates between five types of access. First, educational access which refers to whether test takers have the opportunity to learn the content and familiarise themselves with the different types of tasks and mental demands. Second, financial access which can be defined as test takers' ability to afford a test. Third, geographical access which refers to test takers' ability to have access to a test site. Fourth, personal access which means that test takers who have physical and / or learning disabilities can have access to tests with appropriate test accommodations. Finally, conditions or equipment access which is defined as test takers' familiarity with the test taking equipment, procedures and conditions.

### **1.5.4 Administration**

According to Kunnan (ibid), this quality has to do with physical conditions, and uniformity or consistency. Physical conditions refer to the appropriate conditions for test administration, such as perfect light and temperature levels, and relevant facilities. Consistency refers to 'uniformity in test administration exactly as required' (Milanovic & Weir, 2001, p. 39) in order to achieve similarity in all test sites. It is related to length, materials and any other conditions, for instance time planning. Thus, all test takers, except those who have physical and / or learning disabilities, can pass the test under the same circumstances. Uniformity also includes test security; that is, a test has to be administered in

secure places.

### **1.5.5 Social consequences**

Kunnan (2000) states that the social consequences of a test are related to washback and remedies. The former can be defined as the effect that a test has on instructional practices, such as teaching, materials, learning and test-taking strategies. The latter refers to the remedies that test takers receive to reverse the harmful consequences of a test, such as re-scoring and re-evaluation of test responses, as well as legal remedies for high-stakes tests. There are two main fairness questions to be asked here which are: 'Are the social consequences of a test and/or the testing practices able to contribute to societal equity or not?' and 'Are there any pernicious effects due to a specific test or testing programme?'

To conclude, these five test fairness qualities, which are validity, absence of bias, access, administration and social consequences, play an important role in test fairness and testing practices when they work all together.

### **1.5.6 Implications for test development**

This model of test fairness has significant implications for test development mainly because the issue of test fairness cannot only be discussed after the test is developed and then administered. Test fairness has to be considered at every stage of test development which are design, development, piloting and administration, and use. In fact, the focus may be on different fairness qualities at various levels (ibid).

However, another solution is to have test developers from different groups with regard to gender, race, native language and age, and to train them in fairness issues before test development. Thus, this could facilitate test developers' understanding of the various aspects of fairness.

Finally, concerning the question of 'Who is responsible for fairness-testing practices. Should it be the test developer or the test user, or both?' Kunnan (ibid) believes that since both test developers and test users are the two primary stakeholders of every test, they should be responsible for promoting fairness.

## **1.6 Language Test Design and Development**

Bachman (1990) claims that operationalising the construct that will be tested is considered as a principal step in the development of language tests. Many test writers tend to use tables of specifications to guarantee that the content of the test reflects the content of the syllabus, or that the various parts of the test are equal in content. Of course, in order to achieve this goal, the table of specifications has to be as detailed as possible. However, a table of specifications of English grammar, for instance, might provide a thorough list of the different patterns and structures to be dealt with on one hand, and may give little details about the test method facets on the other hand. Therefore, it is said that tables of specifications succeed in providing a lot of details about the components of ability to be measured; however, they often fail to give precise information about the facets of the test tasks.

Proponents of criterion-referenced testing have always supported the idea of including detailed descriptions or procedures that generate individual test items into the traditional tables of specifications. As a result, approaches, such as 'item forms' (Hively et al., 1973), 'simplified objectives' (Popham, 1974) and 'test specification' (Popham, 1984), have been widely used by test writers. These approaches not only define the ability to be tested, but also describe the attributes of both stimulus and response. In 1988, Griffin et al. applied this approach to the development of language test and found that the detailed descriptions of both item characteristics, as well as its objective can create a close link between patterns of test performance and teaching and learning sequences in language programs (cited in Bachman, 1990).

With reference to all what has been discussed before, we believe that the currently used tests are not reliable and valid assessment tools. Because they do not assess learners' abilities, it can be said that they do not yield useful tests scores that reflect learners' achievement and progress in the English language.

## **Conclusion**

As a conclusion to this chapter, one can say that language assessment is a valuable and integral part of instruction. Well-designed assessment procedures are vital to meet EFL students' needs and improve instruction. It was also shown that language tests, which are



one type of assessment tools, can have potential impact on students' learning if they are well-devised and developed. Moreover, there are two main concepts which are related to test development and test evaluation which are test usefulness and test fairness. These concepts include different qualities or components which must be essentially considered when developing language tests.

In the next chapter, discussion will focus on the speaking skills and the ways used to teach and assess them.

**CHAPTER TWO :**

**THE SPEAKING SKILL : FUNDAMENTALS**

## **CHAPTER TWO: THE SPEAKING SKILL: FUNDAMENTALS**

Introduction	p 28
2.1 Definition of Speaking	p 28
2.2 Components Underlying Speaking Effectiveness	p 30
2.2.1 Grammatical competence	p 30
2.2.2 Discourse competence	p 30
2.2.3 Sociolinguistic competence	p 30
2.2.4 Strategic competence	p 31
2.3 Teaching Speaking	p 31
2.3.1 Interaction as the key to improving EF L learners' speaking abilities	p 31
2.3.1.1 Small talk	p 31
2.3.1.2 Interactive activities	p 32
2.3.1.2.1 Aural: Oral activities	P 32
2.3.1.2.2 Visual: Oral activities.	p 32
2.3.1.2.3 Material-aided: Oral activities	p 33
2.3.1.2.4 Culture awareness: Oral activities	p 33
2.3.2 A Model for teaching speaking	p 34
2.3.2.1 Stage 1: Focus learners' attention on speaking	p 36
2.3.2.2 Stage 2: Provide input and / or guide planning	p 36
2.3.2.3 Stage 3: Focus learners' attention on speaking	p 37
2.3.2.4 Stage 4: Focus on language /skills /strategies	p 37
2.3.2.5 Stage 5: Repeat speaking tasks	p 37
2.3.2.6 Stage 6: Direct learners' reflection on learning	p 38
2.3.2.7 Stage 7: Facilitate feedback on learning	p 38
2.4 Assessing Speaking	P 39
2.4.1 The importance of rating in the speaking assessment cycle	P 40
2.4.2 Designing the rating process	p 40
2.4.3 Useful speaking scores	p 41
2.4.4 Speaking assessment procedures	p 41
2.4.4.1 Communication-oriented tasks and criteria	p 41
2.4.4.2 Pair and group tasks	P 42
2.4.4.3 Rating checklist	P 42
Conclusion	P 43

## **CHAPTER TWO: THE SPEAKING SKILL : FUNDAMENTALS**

### **Introduction**

This chapter discusses the importance of the speaking skill and the ways used to teach and assess it. First, it provides some definitions of speaking and introduces the term of the construct. Then, it states the components underlying speaking effectiveness and reviews the communicative competence proposed by Canale and Swain. Next, it tackles the difficulties and challenges facing teachers in teaching this basic skill, and how interaction plays an important role in enhancing EFL learners' speaking abilities. Also, this chapter displays a model for teaching speaking which is constituted of a series of seven complementary stages that can help improve learners' speaking competence. Finally, the chapter discusses speaking assessments and the procedures used to rate performances.

### **2.1 Definition of Speaking**

Speaking is an integral part of people's daily lives. Bygate (1987) illustrates that people are very often judged by what they say, and that through their speech they can make or lose friends. Thus, speaking is the skill that serves as a means of social solidarity, social ranking, professional development and business. Besides, oral communication is a reciprocal process which involves at least two people, namely speaker and listener. It requires the productive skill of speaking and the receptive skill of listening. While listening refers to the process of decoding or interpreting a message, speaking is defined as the process of encoding the message the speaker intends to convey using appropriate language.

Flucher (2014) argues that defining speaking or identifying the constituents of speaking is basically a question of defining construct. However, before providing definitions for the construct of speaking, there has to be a clear understanding of the term construct. It is often stated that a construct is a concept that is 'deliberately defined for special scientific purpose'. According to Kerlinger and Lee (2000), a construct is different from a concept due to two main reasons. First, it can be observed and measured. Second, the relationship which exists between various constructs constitutes a theory (cited in *ibid*). Hence, if speaking is to be defined as a construct, it has to be related to observable 'things' that can be graded. Then, the relationship between 'speaking' and 'listening' or 'reading', for instance, may establish a theory of 'language use'. Such definitions are considered to be

operational.

However, Lado (1961) sees that the construct of speaking could not have an operational definition. As a result, he suggests testing linguistic elements to avoid confusions between the construct and variables, such as 'talkativeness' and 'introversion'; that is to have a language test as a test of speaking in order to exclude anything that may negatively influence the construct and the interpretation of the scores (cited in *ibid*). This approach is known as the 'Trait Theory' approach to construct validity. In 1999, Chapelle argued that it may be more convenient to consider the contextual factors when defining the construct since they can affect discourse and test score, as it is the case for tests of speaking for specific purposes (cited in *ibid*). In these cases, the construct definition, task choices and scoring procedures are designed to obtain evidence for the speaking ability in a small subset of contexts. Therefore, the definition of the construct, its range and generalisability should be based on test purpose.

However, the definition of construct could include many other factors. For example, it may consider aspects of talk, such as the degree of interlocutors support if it is accepted that an act of conversation is created between interlocutors talking in particular situation.

In short, there is no operational construct definition that can capture all the elements or aspects of human communication which is characterised by its complexity. Also, there is no test of speaking that measures everything; consequently, selections are based on test purpose and the interpretations of the scores.

Our knowledge about what is meant by to 'speak' can be used to define constructs that are useful for the different testing purposes. The construct definition should then be judged according to its utility in making inferences from test scores, and its usefulness for establishing a validity argument that creates a link between a test score and its interpretation.

Thus, a construct can be made using research findings about speaking and testing second language teaching and through a process of 'pick and mix'; that is, there has to be a rationale and empirical evidence to support the 'mix' in terms of test purpose. However, Butler et al. (2000, p.3) states that 'the available research does not provide a firm foundation for constructing a specific test of speaking as part of second language academic communication competence.' (cited in *ibid*). This means that there is no true construct definition for second language academic communicative competence and there is still a need for more research.

## **2.2 Components Underlying Speaking Effectiveness**

Learning to speak a foreign language is not limited to learning its grammatical and semantic rules. EFL learners must also acquire knowledge about how native speakers use language appropriately in different situations.

Canale and Swain (1980) propose that communicative competence consists of grammatical competence, discourse competence, sociolinguistic competence and strategic competence (cited in Richards & Renandya, 2002).

### **2.2.1 Grammatical competence**

This concept is defined as the knowledge of morphology, syntax, vocabulary and mechanics. With reference to speaking, the term mechanics refers to basic sounds of letters and syllables, pronunciation of words, stress and intonation. Thus, grammatical competence contributes to EFL learners' fluency as it enables them to accurately and effectively use and understand English language structures (ibid).

### **2.2.2. Discourse competence**

Discourse competence is related to intersentential relationships. In other words, it refers to the knowledge of the rules of cohesion and coherence which serve as a means to maintain communication together in a meaningful way. In an act of communication, the interlocutors have to comprehend and produce language. This requires the ability to perceive and process sequences of discourse and to formulate utterances which are in accordance with the previous and the following utterances. That is, effective speakers must have a large knowledge of the structures and discourse markers to express ideas and show different relationships, such as cause, emphasis and contrast. Consequently, developing discourse competence can help learners successfully perform turn-taking in a conversation (ibid).

### **2.2.3 Sociolinguistic competence**

Knowledge of language does not alone enable learners to effectively and appreciately use the foreign language being learnt. They must also know what is socially and culturally expected of them. In other words, learners must also learn the social and cultural rules and

norms of the target language as they help them know when and how to perform the different speech acts. For example, learners will be able to know how to ask questions, what comment are appropriate for particular contexts, as well as how to respond nonverbally with regard to the purpose of the talk (ibid).

#### **2.2.4 Strategic competence**

Shumin (2002) claims that strategic competence is perhaps the most important component of communicative competence as it is 'the way we manipulate language in order to meet communicative goals' (Brown, 2000, p. 248). That is, it is the capacity to compensate for deficient knowledge of linguistic, sociolinguistic and discourse rules. With regard to speaking, strategic competence refers to the knowledge of the different mechanisms of turn-taking and their realisation. In other words, it is the ability to know when and how to take or give the floor, how to keep the conversation going, how to end a turn or a conversation, and how to repair communication break down and comprehension problems.

### **2.3 Teaching Speaking**

Speaking is one of the basic elements of communication; as a result, it requires more and special attention in EFL teaching. Teachers have to consider the different components of communicative competence which underlie speaking effectiveness as well as the various factors and conditions that can affect oral communication. In addition, teaching speaking effectively involves sufficient language input and speak-promotion activities which help EFL learners improve their speaking performance.

#### **2.3.1 Interaction as the key to improving EF L learners' speaking abilities**

##### **2.3.1.1 Small talk**

Developing EFL learners' ability to carry out real conversations with people in society can be through engaging them in brief, casual conversations about different topics, such as weather, sports, vocations and food. In other words, EFL learners should first develop speaking skills in short, interactional exchanges which involves making only one or two utterances at a time (Richards & Renandya, 2002).

### **2.3.1.2 Interactive activities**

Most EFL learners practise the target language only in the classroom. So, they should be given enough opportunities to speak and interact. Teachers have to motivate their learners to communicate; in other words, they have to provide learners with a reason or need to speak. For example, teachers can stimulate EFL learners to talk by extensively exposing them to authentic language using audio-visual aids and providing them with opportunities to use that language.

Nunan (1989) argues that since most real-life situations involve more than one macroskill, teachers should consider all the skills and they should integrate them when designing interactive activities. He argues that interactive activities can be characterised as being effective only when they are manipulative, meaningful and communicative. That is, they should be based on naturalistic or authentic materials. Also, they should allow learners to use and practise specific language features and communicative skills which they need them in real life. Finally, they should stimulate and activate learners' psycholinguistic processes. With reference to these criteria, there are four types of activities which are considered to be appropriate to eliciting oral production because they enable learners to learn from visual and auditory experiences and to develop flexibility in their learning styles. Also, they require the use of different learning strategies and behaviours for different tasks (Richards & Renandya, 2002).

#### **2.3.1.2.1 Aural: Oral activities**

If aural activities, such as news reports on the radio, are well-selected and adapted to the level of particular groups of learners, they can be used in some productive activities as background or as input for interaction. For example, teachers can ask learners to listen to tape recordings and then to act them out. Also, teachers can divide students into small groups and have each group listen to a different segment of a story. Then, students are provided with a work a worksheet of comprehension questions and work together on information-gap activity. Hence, this motivate students to speak. This technique is called jigsaw listening.

#### **2.3.1.2.2 Visual: Oral activities.**

It is essential for learners to be exposed to many situations, accents and voices as they may have few apportunities to interact with native speakers. Thus, audio-visual materials,



such as films, videotapes and soap operas, can be used to meet this need. Carrasquillo (1994) reports that the audio-visual materials can motivate learners and expose them to a broad range of authentic speech with different registers, accents, intonation, rhythms and stresses, as well as to language used in real situations (cited in *ibid*). The audio-visual aids enable students to learn the different levels of formality as well as the nonverbal behaviour and types of exclamations and fillers which are used. They can also observe how people start, maintain and end a conversation. What deepens students' learning is subsequent practice of dialogues, role-playing and dramatisations.

Visual materials can be used in different ways as a starter for interaction. For example, the story of a movie can be used to elicit opinion-expressing activity. Similarly, teachers can play nonverbal videos and then ask students to describe what they have viewed. Students can focus on the content and imitate the 'model's' body language while they are watching the videos. Therefore, students will gradually assimilate the verbal and nonverbal messages. They will also be able to communicate naturally.

#### **2.3.1.2.3 Material-aided: Oral activities**

Appropriate reading materials which are adapted to the level of students and accompanied by comprehension questions can result in a creative production in speech. Series of pictures or cartoon strips can also be used to stimulate storytelling, and articles in newspapers or from well-designed textbooks can lead to the production of oral reports or summaries. Hotel brochures can be used for making reservations, and menus can be used for ordering in a restaurant or making purchases in the supermarket. There is actually a wide range of sources that can serve as language input for oral activities and form a basis for communicative tasks which prepare learners for real life situations (*ibid*).

#### **2.3.1.2.4 Culture awareness: Oral activities**

Culture plays an influential role in shaping speakers' communicative competence which is related to the appropriate use of language. Appropriateness is generally defined in terms of the shared social and cultural conventions of a particular speech community. Thus, it is very important to learn the socio-cultural rules and norms of the foreign language in order to be able to communicate effectively and appropriately. A wide range of cultural assumptions need to be taught explicitly to speakers from other cultures as they are

normally presupposed and not made clear by native speakers (Brown & Yule, 1983). Therefore, cultural learning which is illustrated by activities and reinforced through practice will motivate learners.

Teachers can present situations in which cultural misunderstandings make people feel offended, angry and confused and then ask students questions and carry out class discussions. Teachers can ask students to analyse and identify the causes of misunderstandings. This will make them think about how native speakers act and perceive things. Hence, students will get deeper insight into the target culture. This kind of exercise can create a balance between students' native culture and the necessity to learn the target culture. At the end, it will improve their culture awareness (Richards & Renandya, 2002).

In general, the use of audio-visual materials brings, sight, hearing and kinesthetic participation into interplay. In this case, students do not just imagine the situations, but they experience them. Also, task-oriented activities can give students a purpose to talk. The flexibility and adaptability of these activities are fundamental in order to meet the communicative needs of learners. It is necessary to follow open language experiences with more intensive structured situations, dialogues and role-plays because the time available in the classroom is limited. Consequently, students will get opportunities to use the target language and to develop their self-confidence.

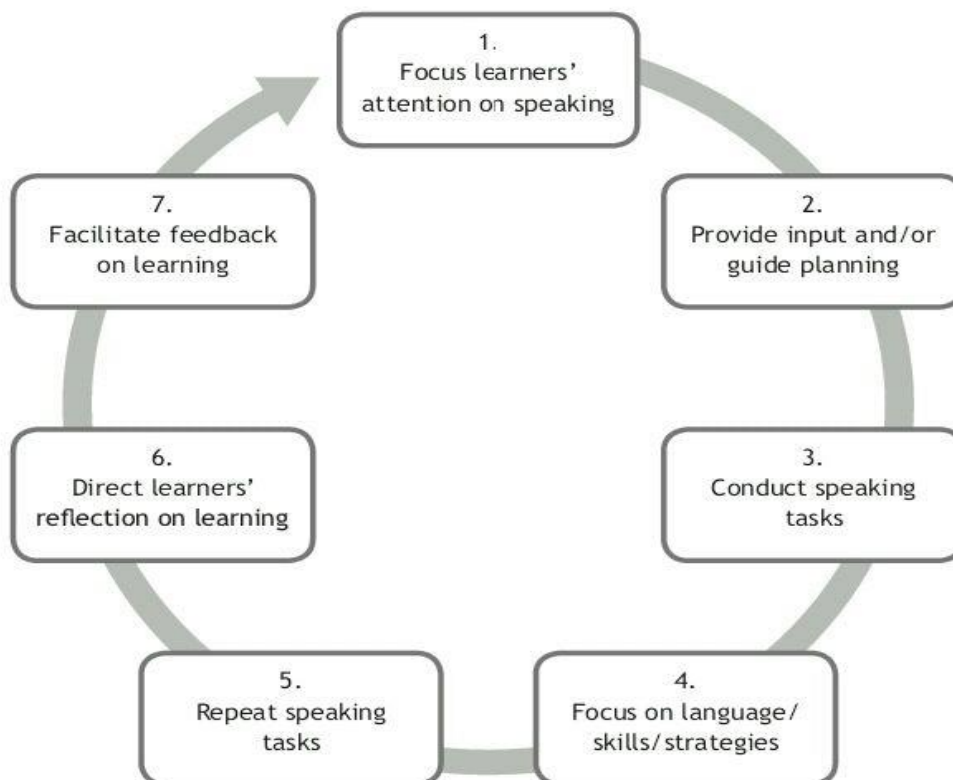
### **2.3.2 A model for teaching speaking**

There is no doubt that speaking is the main language skill. However, when it comes to teaching speaking, many language teachers face many difficulties and challenges in an attempt to enhance their students' speaking performance. Goh and Burns (2012) believe that despite the fact that many language teacher make great efforts when designing appropriate activities for students to practise their speaking in classrooms, many of these activities only establish contexts and provide opportunities for oral interaction. In other words, if learners are talking to each other in the target language, this does not mean necessarily that they are improving their speaking skills. It is said that learners develop their speaking abilities randomly or incidentally because, unlike teaching the writing and reading skills, teachers do not have a model which guides them when teaching speaking.

Therefore, Goh and Burns (ibid) suggest that teachers can help learners develop their speaking competence if they encourage to do the following:

- Use a wide range of core speaking skills.
- Develop fluency in expression of meaning.
- Use grammar flexibly to produce a wide range of utterances that can express meaning precisely.
- Use appropriate vocabulary and accurate language forms relevant to their speaking needs.
- Understand and use social and linguistic conventions of speech for various contexts
- Employ appropriate oral communication and discourse strategies.
- Increase awareness of genre and genre structures.
- Increase their metacognitive awareness about L2 speaking.
- Manage and self-regulate their own speaking development. (pp. 151–152)

Many speaking activities that teachers use in their classrooms are considered to be temporary and disconnected. As learners do not often record the outcomes of such activities, it is difficult for learners to recall or understand their purposes. Also, learners are often unable to state what they have learnt. As a result, Goh and Burns (*ibid*) presented a teaching cycle which can guide the planning and sequencing of the learning activities and which aims at developing students' speaking competence.



**Figure 2.1** The Teaching-speaking cycle (*ibid*, p. 153)

Figure 2.1 presents the model for teaching speaking. It shows a sequence of learning activities in the form of a cycle. At each stage, the role of the teacher is essential as a facilitator of learning and practice and as a provider of input and feedback. Also, in order to make learners get many opportunities to speak, as well as to work together, the various stages of the cycle includes dialogue and collaboration among peers.

Although every stage in the teaching cycle serves the broad developmental objectives for speaking, some stages can support more than one objective. By choosing and planning different activities, the teacher can decide on which objective to focus on in a given cycle.

### **2.3.2.1 Stage 1: Focus learners' attention on speaking**

This first stage is related to raising learners' metacognitive activities about speaking a foreign language. These metacognitive activities focus on the three types of metacognitive knowledge which are person knowledge, task knowledge and strategic knowledge. This stage has two main purposes.

a) to encourage learners to plan for overall speaking development

Learners are given prompts to make them think about the demands of speaking a foreign language and how they can prepare themselves for it.

b) to prepare learners to approach a specific speaking task

Here, the prompts focus on the speaking task which has been designed for the teaching cycle and are used to prepare learners as they familiarise themselves with the outcomes of the task and as they consider the strategies they need to complete it. The prompts can also serve as stimuli for learners' knowledge about the requirements of the task (ibid)

### **2.3.2.2 Stage 2: Provide input and / or guide planning**

Speaking a foreign language is considered to be a challenge as it can cause anxiety to many language learners. Hence, a teacher should provide support to learners by giving them enough time to plan what to say and how to say it. In other words, this stage has to do with scaffolding learning which Mercer and Steiter (cited in Goh & Burns, *ibid*) define as the help provided to learners in order to enable them to accomplish a task which they would not do on their own, and which aims at bringing learners closer to the state of competence so that they will be able to complete such kinds of tasks without assistance. This preparation stage has the following purposes:

- introducing or teaching new language
  - enabling learners to reorganise their developing linguistic knowledge
  - activating existing linguistic knowledge
  - recycling specific language items, and easing processing load
  - pushing learners to interpret tasks in more demanding ways
- (Skehan, 1998, pp. 137–139)

### **2.3.2.3 Stage 3: Focus learners' attention on speaking**

The purpose of the third stage is to establish a context where learners can practise speaking through a communicative task which encourages them to express meaning using the linguistic knowledge, skills and strategies which they have. That is, at this stage, learners are encouraged to develop fluency without paying too much attention to accuracy. Thus, This stage is less demanding for the learners as they are guided by the teacher or individual pre-task planning that has occurred in the second stage (Goh & Burns (2012).

### **2.3.2.4 Stage 4: Focus on language /skills /strategies**

The purpose of this stage is to create opportunities for learners to enhance language accuracy, as well as their effective use of skills and strategies. The teacher has to draw learners' attention to some parts of the fluency task they have completed. These parts are related to language features, such as vocabulary, grammar, pronunciation and text structures (ibid).

### **2.3.2.5 Stage 5: Repeat speaking tasks**

At this stage, learners perform the speaking task(s) from Stage 3 again. However, learners have now the chance to analyse and practise selected language items or skills during Stage 4 in order to improve their speaking performance. Repetitions could be carried out by repeating parts of the original task or the entire task, or by having students change groups or partners. Also, teachers can introduce a new task similar to the one learners have just done, for example, instead of giving instructions about making your favourite food, learners could give instructions about a topic of their own choice, so that they rehearse a procedure genre again (ibid).

### **2.3.2.6 Stage 6: Direct learners' reflection on learning**

This stage aims at encouraging learners to self-regulate their learning through monitoring and evaluating what they have learnt from the preceding stages. Reflection can be done individually, in pairs, or even in small groups. Individual and group reflection often has a cathartic effect on learners who may feel stressed and anxious, and who think that they are the only ones feeling that way. Teachers can also encourage their learners to

start their own written or audio blogs in order to record their experiences about learning to speak a foreign language.

Learners' reflection should be guided by different types of metacognitive knowledge, and can focus on one or more of the following points:

- demands of the speaking tasks which they have become aware of;
- the strategies that are useful to meet the demands of the task;
- their informal assessment of their capabilities and performance;
- areas of their performance that show improvement; and
- areas to be further improved.
- plans for improving specific areas (ibid).

### **2.3.2.7 Stage 7: Facilitate feedback on learning**

At the last stage of the Teaching Speaking Cycle, the teacher provides learners with feedback on their performance in earlier stages. This feedback can take many forms, for example, comments or grades on an individual learner's skills and performance from observation sheets used during the speaking task, or written comments in learners' journals or blogs. Also, learners can exchange their written individual reflections and comments on each other's progress and achievements. In addition, the teacher can give consolidated comments based on written reflections from the class (ibid).

The cycle that has been presented is not meant to be completed in just one or two lessons. However, it is an overall approach to supporting and scaffolding the learning of speaking skills and strategies that can be introduced and extended over several lessons or even a unit of work. For example, a cycle can consist of a series of lessons based on a specific theme or topic.

What is interesting about this cycle is that it helps teachers plan systematically each component and organise relevant tasks and materials for learners at different stages of learning. It also considers activities that engage learners cognitively and affectively. On the whole, this cycle emphasises five main points. First, the teaching of speaking should give the most importance to the roles played by the teacher, the learner and the materials. Second, the main aim of speaking tasks is to help students develop fluency; that is, it is to communicate meaning appropriately with few hesitations. This can be achieved through the use of accurate language and discourse routines, appropriate speaking skills, and

effective communication strategies.

Third, learners' speaking performance can be enhanced through pre-task planning and task repetition because these activities can reduce cognitive load during speech processing. Fourth, as learning requires noticing key information and storing it in the long term memory, activities that draw learners' attention to language, skills and strategies are very important. Finally, activities that help learners develop metacognitive knowledge and self-regulation are also needed to address affective and other cognitive demands of learning to speak a foreign language.

Therefore, by planning lessons according to the Teaching Speaking Cycle, teachers can address all these concerns and provide valuable scaffolding for learners as they engage in speaking tasks. Learners will not only practise expressing meaning using their existing language resources, but they will also receive timely input and guidance for improving their performance.

## **2.4 Assessing Speaking**

Speaking is an interactive process of constructing meaning that involves producing, receiving and processing information (Florez, 1999) (cited in Comings et al., 2006, p. 124). Thus, because of its interactive nature, speaking is often tested in live interaction. However, assessing speaking requires special procedures in order to ensure the reliability and validity of the scores. One of the reasons is that the test discourse is completely unpredictable. Although the speakers in a test have to talk about the same topic and have the same roles and aims, it is impossible to find two conversations that are exactly the same. Also, as the rating process involves human raters, there is some inevitable variability.

### **2.4.1 The importance of rating in the speaking assessment cycle**

The process of developing and using speaking tests is considered to be cyclical as it starts with a purpose and finishes with the interpretation of the scores for that purpose. What goes between these two stages is the process of test development and also a two-step testing process which includes test administration and scoring. The former refers to the interaction between the test takers and the interlocutors or other test takers to accomplish the speaking task, and results in the test discourse. The latter refers to the interaction which

occurs between the raters, the criteria and the performances in order to produce scores or grades. Because the test takers' experience of the test is built on the tasks, the focus in designing speaking assessments is on them. However, the validity of the scores depends much more on the rating criteria and their relationship to the task. Thus, Grove and Yule (2001) state that the criteria must not be developed in isolation of the tasks, otherwise there is a risk of mismatch which causes the loss of information about the quality of performance (cited in Luoma, 2004). However, it is important to draw attention to the fact that in order to put the criteria into practice the rating process has also to be planned carefully in order to obtain meaningful scores.

#### **2.4.2 Designing the rating process**

The rating process shows exactly how the criteria will be used to rate performances. It is based on the structure of the test and the definition of the test construct. The short responses of taped-based and highly structured live tests are scored one by one and then the points are added up into task or section scores and total scores. Meanwhile, live tests are often scored task by task. However, because of the length of performances, it is necessary to use scales instead of points. Hence, the criteria used to score a monologic description task might be comprehensibility, discourse organisation, accuracy and fluency; whereas, a discussion task might be scored for interaction skills, appropriateness and vocabulary range. Then, an overall grade is usually obtained by averaging the task scores. The raters might also be asked to listen to the entire performances and give overall scores. But, in this case they are required to provide detailed ratings which may be reported to the examinees as feedback.

Test developers also have to design a rating form for their test in order to gather the detailed rating information. Rating forms are the concrete outcome of the design of the rating process. They are used by raters to record their ratings. They are so useful as they contribute to the arrangement of the process, its speed, and its consistency (ibid).

#### **2.4.3 Useful speaking scores**

Speaking scores, like the scores of any type of tests, must be fair, dependable and most importantly useful for the intended purposes. Technically speaking, they must be reliable and valid. Reliability is usually defined as score consistency (AERA, 1999; Brown and



Hudson, 2002; Hennig, 1987) (cited in *ibid*). In other words, the scores of a test can be reliable only if they do not change almost completely if the same test is given to the same group of test takers at another time. Thus, reliability is an essential quality in order to make any decision as it means that the scores are dependable. Concerning validity, it is defined as the meaningfulness of scores. According to Bingham (1937), validity is 'the correlation of scores on a test with some other objective measure of that which it correlates' (cited in Wainer & Braun, 2009, p. 40). It is an important concern in test development. It is related to content coverage and comprehensiveness in accordance with the test purpose, as well as to the match between the test and the non-test activities which the scores are expected to show. It is also related to the effects the test has on both score users and examinees. Moreover, speaking test has to be practical; that is, there must be enough source; otherwise, it will not be developed and arranged. The main qualities of language tests are all discussed under the concept of usefulness (Bachman & Palmer, 1996) in chapter 1.

#### **2.4.4 Speaking assessment procedures**

Speaking assessment requires consideration of many elements, such as the choice speaking-assessment tasks and overall assessment procedure, to ensure the reliability and validity of test scores.

##### **2.4.4.1 Communication-oriented tasks and criteria**

Communication-oriented tasks have been used in speaking test for years to measure test takers' skills to communicate effectively and appropriately. However, communication-oriented criteria have recently appeared. Test developers have to identify the parameters that control learners' success by observing their speaking performances in order to decide on the appropriate communication-oriented criteria. For example, in information-related talk, such as explanation or narratives, a test taker's success is partly based on the content and their abilities to sequence their talks. Another part of a speaking test may depend on other types of criteria, such as fluency, accuracy, the speaking strategies and how well the test takers take the listener into account (Luoma, 2004).

#### **2.4.4.2 Pair and group tasks**

As speaking assessments are time-consuming, the administration may be speeded up by using tape-based tests. However, this does not really solve the problem because the raters still have to listen to the entire performances. Therefore, they can make two or more test takers interact with each other in order to save testing and scoring time in live tests. This way, peer interaction is considered to be more realistic than the kind of interaction which is between the interlocutor and one test taker. However, pair and group tasks have also some disadvantages. It is argued that the developers have the whole responsibility for showing their own and their peers' language skills. As a result, there has to be some ways to make pair and group tasks more effective. This can only be achieved by observing and analysing learners' performances to discover the strengths and weaknesses of different tasks. For example, there have to be efficient ways of giving task instructions and presenting task materials that support the test discourse rather than make it artificial. Also, many studies concerning the appropriateness of information gap tasks for test takers of different ages and abilities will need to be conducted. Moreover, the different picture-based or text-based materials will need to be studied in order to investigate their benefits and drawbacks (ibid).

#### **2.4.4.3 Rating checklist**

Rating checklists are a useful assessment tool as they help raters note their comments quickly, and they also provide test takers with interesting or valuable information about their performances (ibid). Checklists are considered to be a 'dichotomous' form of rating as completing a checklist means making judgements about the presence or the absence of the specific behaviours or skills which are listed (Brassard & Boehm, 2007).

By and large, although speaking assessments are necessary and important, they are generally time-consuming and require a lot of efforts. However, the development of better speaking assessments can be achieved by considering their three main stages, which are the development of tests, their administration and their use. In other words, the practices of speaking assessments can be improved through 'action, reflection and reporting' (Luoma, 2004); that is, by observing and analysing test takers' performances, as well as studying the different types of tasks that are used in tests and investigating the best ways to avoid or at least reduce their drawbacks. Also, testers' experiences can help learn more about this field and increase our knowledge.

With regard to what has been mentioned before, we believe that group tasks or group assessment can be more appropriate if we adopt speaking tests in Algeria. This is because of several reasons. For example, since classes in Algeria are overcrowded, we think that group assessment can be time saving. Besides, it allows interaction between test takers. This can create real contexts, and enables raters to assess test takers' communication skills.

Moreover, the use of rating checklists can be very useful because they can help raters write their comments quickly. They can also provide test takers with information about their weaknesses and strengths. Hence, they can find ways to improve their speaking performances.

Therefore, we chose group format to test the speaking abilities of the participants in the present research. We will question them about the effectiveness of group assessment, and whether it create a comfortable context for speaking assessment.

## **Conclusion**

Throughout this chapter, we have discussed the fact that speaking is a challenging skill for EFL learners as well as teachers as it requires a great deal of efforts to teach and assess it. It has been shown that learners should have enough opportunities to practise speaking in classroom in order to enhance their speaking abilities. Also, the process of teaching speaking should be cyclical. In other words, teachers should not give their students separate and short-term activities. There should be a series of interrelated stages where the focus is on both the learner and the teacher as a provider of input and feedback. Finally, educators and teachers must not neglect the process of speaking assessment as it is an indicator of learner progress and achievement; consequently, it requires more attention.

The following chapter will be devoted to the research methodology underlying this study. Furthermore, it will provide the interpretation of the collected data followed by a discussion and synthesis of the major findings.

**CHAPTER THREE :**  
**FIELDWORK AND DATA ANALYSIS**

## **CHAPTER THREE: FIELDWORK AND DATA ANALYSIS**

Introduction	p 46
3.1 Research Methodology for this Dissertation: Theoretical Background	p 46
3.1.1 Research approach	p 46
3.1.1.1 Qualitative approach	p 47
3.1.1.2 Quantitative approach	p 47
3.1.1.3 A Mixed-methods approach	p 48
3.1.2 Research design / strategy	p 49
3.1.2.1 Qualitative designs	p 49
3.1.2.1.1 Narrative	p 50
3.1.2.1.2 Phenomenology	p 50
3.1.2.1.3 Ethnography	p 50
3.1.2.1.4 Case study	p 51
3.1.2.1.5 Grounded theory	p 51
3.1.2.2 Quantitative Designs	p 51
3.1.2.2.1 Surveys	p 51
3.1.2.2.2 Experiments	p 52
3.1.2.3 A mixed-methods design	p 52
3.1.2.3.1 Convergent parallel-mixed- methods design	p 52
3.1.2.3.2 Explanatory sequential-mixed- methods design	p 53
3.1.2.3.3 Exploratory sequential-mixed- methods design	p 53
3.1.3 Data collection methods	p 53
3.1.3.1 Qualitative data collection methods	p 54
3.1.3.1.1 Observation	p 54
3.1.3.1.2 Interviews	p 54
3.1.3.1.3 Questionnaires	p 55
3.1.3.1.4 Focus group	p 55
3.1.3.2 Quantitative data collection methods	p 55
3.1.3.2.1 Tests	p 55
3.1.3.2.2 Structured questionnaires	p 56
3.1.3.3 A Mixed-methods data collection methods	p 56
3.1.4 Sampling	p 57
3.2 Research Methodology for This Dissertation: Procedures	p 58

3.2.1 Research approach	p 58
3.2.2 Research design / strategy	p 58
3.2.3 Data collection methods	p 58
3.2.3.1 Questionnaires	p 59
3.2.3.1.1 Aim	p 59
3.2.3.1.2 Structure and content	p 59
3.2.3.1.3 Piloting and validation	p 60
3.2.3.2 Interview	p 61
3.2.3.2.1 Aim	p 61
3.2.3.2.2 Structure and content	p 61
3.2.4 Data collection procedures	p 61
3.2.4.1 Data collection proceduees for the questionnaires	p 61
3.2.4.2 Data collection proceduees for the interview	p 62
3.2.5 Population and sampling technique	p 62
3.2.6 Data analysis procedures	p 63
3.2.6.1 Data analysis procedures for the questionnaires	p 63
3.2.6.2 Data analysis procedures for the interview	p 63
3.3 Results	p 64
3.3.1 Pre-questionnaire	p 64
3.3.2 Post-questionnaire	p 72
3.3.3 Interview	p 78
3.4 Discussion of the Results	p 81
3.5 Systhesis of the Findings	p 84
Conclusion	p 85

## **CHAPTER THREE: FIELDWORK AND DATA ANALYSIS**

### **Introduction**

The present chapter is devoted to the fieldwork and the analysis of the collected data. First, it reviews the research methodology which is used in the present study to investigate learners' attitudes towards the implementation of a Formal Oral Test. It describes the adopted approach, the strategy, and the methods employed to collect the data, and it provides a justification for each selection. Then, it displays a full data analysis followed by a discussion based on the obtained results. This chapter ends with some recommendations in order to improve English language learning and assessment.

### **3.1 Research Methodology for this Dissertation: Theoretical Background**

With regard to the issue of research methodology, we believe that it is the central element in any study. It can be defined as a way to systematically solve the research problems. 'In it we study the various steps that are generally adopted by a researcher in studying his research problem along with the logic behind them' (Kumar, 2008, p. 5).

Thus, we find that the conceptual framework proposed by Hoadjli (2016), which is based on two other conceptual frameworks suggested by Creswell and Denscombe (2014), sounds interesting since it suits our research study. The main components of this framework are:

- 1- Research approach,
- 2- Research strategies, and
- 3- Data collection methods.

#### **3.1.1 Research approach**

'Research approaches are plans and the procedures for research that span the steps from broad assumptions to detailed methods of data collection, analysis, and interpretation' (Creswell, 2014, p. 3). There are three research approaches which are qualitative, quantitative and mixed- methods. According to Newman and Benz (1998), qualitative and quantitative approaches represent different ends on a continuum. Thus, a mixed-methods

study exists in the middle of this continuum as it combines qualitative and quantitative elements.

### **3.1.1.1 Qualitative approach**

A qualitative approach, as asserted by Nunan (1992), uses textual analysis and is very effective in the exploration and interpretation of participants' beliefs, opinions, attitudes and motivation. A similar definition states that 'qualitative approach is concerned with subjective assessments of attitudes, opinions and behaviour' (Kothari, 2004, p.5).

According to Hoadjli (2016), the main characteristics of a qualitative research are:

- Emergent research design: The emergent research design means a qualitative research study is kept open and fluid so that it can respond in a flexible way to new details or opinions that may emerge during the process of investigation.
- The nature of qualitative data: Qualitative research works with a wide range of data, including interviews, documents, and even images.
- The characteristics of the research setting: Because of the qualitative approach nature that seeks to describe social phenomena as they occur naturally, qualitative research takes place in natural setting, without any attempts to manipulate the situation under study.
- Insider meaning: Qualitative research is concerned with subjective opinions, experiences, and feelings of individuals. In this respect, the goal of qualitative research is to explore the participants' views of the situation being studied.
- Small sample size: Qualitative research typically relies on the necessity to use smaller samples of participants.
- Interpretive analysis: Qualitative research is fundamentally interpretive, which means that the research outcome is ultimately the product of the researcher's subjective interpretation of data (pp. 33-34).

### **3.1.1.2 Quantitative approach**

According to Neuman (2014), a quantitative study depends on positivistic principles. That is, the focus is on accurately measuring variables and testing hypotheses. Therefore,



a quantitative research is an empirical research where the data are in the form of numbers (Punch, 2005, p. 3). Davies (1995) emphasises that the distinction between qualitative and quantitative research is not only in the fact that the former uses words whereas the latter uses figures. However, the difference lies in the general orientation underlying the study, the method of data collection applied, the nature of the collected data, and the method of data analysis used to process the data and to obtain results (cited in Dörnyei, 2007). Hence, Dörnyei (ibid) reports that ' a quantitative research involves data collection procedures that result primarily in numerical data which are then analysed primarily by statistical methods. Typical example can be a survey research using a questionnaire, analysed by a statistical software such as SPSS' (p. 24).

Therefore, the differences between a qualitative and quantitative research can be summarised as follows:

- quantitative research uses numbers as the unit of analysis;
- qualitative research uses words or visual images as the unit of analysis.
- quantitative research tends to be associated with researcher detachments;
- qualitative research tends to be associated with researcher involvement.
- quantitative research tends to be associated with large-scale studies;
- qualitative research tends to be associated with small-scale studies.
- quantitative research tends to be associated with analysing specific variables;
- qualitative research tends to be associated with a holistic perspective.
- quantitative research tends to be associated with data analysis after data collection; and
- qualitative research tends to be associated with data analysis during data collection (Denscombe, 2014, pp. 245-246).

### **3.1.1.3 A Mixed-methods approach**

Dörnyei (2007) reports that 'a mixed-methods research involves different combinations of qualitative and quantitative research either at the data collection or at the analysis levels' (p. 24). A similar definition states that 'a mixed methods research is an approach of inquiry involving collecting both quantitative and qualitative data, integrating the two forms of data, and using distinct designs that may involves philosophical assumptions and theoretical frameworks' (Creswell, 2014, p. 4). Researchers may use a mixed-methods research because they believe that incorporating elements of both qualitative and

quantitative approaches can result in a more complete understanding of a research problem. However, although both types of data are important for the study, one approach might be more emphasised than the other.

The main characteristics of a mixed-methods research are as follows:

- Expanding the understanding of a complex issue: A mixed-methods approach would broaden the scope of the investigation and enrich the researchers' ability to draw conclusions about the problem under study.
- Corroborating findings through 'triangulation': The use of different data sources, investigators, theories, or research methods generates multiple perspectives on a phenomenon. Such an effective strategy ensures research validity.
- Reaching multiple audiences: Because of the combination of the methods in mixed-methods research, the final results can be more palatable for certain audiences than outcomes of a monomethod study (Hoadjli, 2016, p. 35).

### **3.1.2 Research design / strategy**

When carrying out a study, a researcher has to make many important decisions. One of these crucial decisions is related to the choice of a research strategy or design. 'A strategy is a plan of action designed to achieve a specific goal' (Denscombe, 2014, p.3).

Denscombe (ibid) argues that a researcher must consider three main questions when choosing the research strategy. First, 'is it suitable?' This means that strategies can only be judged on the basis of the purpose for which they are used. In other words, it is better to think of strategies in terms of how useful and appropriate they are. Second, 'Is it feasible?' That is, a researcher must also consider certain practical aspects of conducting research. These aspects are access to data sources, deadlines or time constraints, and the audience for the research and the particular research community within which the researcher is working. Finally, 'Is it ethical?' This means that a researcher must avoid any strategy that may cause harm to the research participants.

#### **3.1.2.1 Qualitative designs**

Despite the fact that there are many qualitative designs, Creswell (2014) recommends qualitative researchers to choose from among the possibilities, such as narrative, phenomenology, ethnography, case study, and grounded theory. Although these five

designs, which are popular across social sciences, fall under the umbrella concept of 'qualitative', each one has somewhat different focus. Therefore, there are some variations in how the research questions may be asked, sample selection, data collection and analysis, and write-up. However, these types of research may also overlap. That is, a researcher may combine two or more, such as in an ethnographic case study (Merriam & Tisdell, 2016).

#### **3.1.2.1.1 Narrative**

A useful definition of it is that offered by Elliott (2005) who proposes that 'a narrative can be understood to organise a sequence of events into a whole so that the significance of each event can be understood through its relation to that whole. In this way a narrative conveys the meaning of events' (p. 3). That is, narrative is the representation of a series of events which are meaningfully connected.

#### **3.1.2.1.2 Phenomenology**

According to Denscombe (2010), 'phenomenology is an approach that focuses on how life is experienced' (p. 94). That is, it is not concerned with the causes of things, but it attempts to describe how things are experienced by the participants. Hence, it deals with:

- people's perceptions or meanings;
- people's attitudes and beliefs; and
- people's feelings and emotion (ibid).

#### **3.1.2.1.3 Ethnography**

Ethnography 'involves direct and sustained contact with human beings, in the context of their daily lives, over a prolonged period of time; draws on a family of methods, usually including participants observation and conversation; respects the complexity of the social world; and, therefore, tells rich, sensitive and credible stories (O'Reilly, 2012, p. 3). A similar definition states that it is the strategy 'in which the researcher is immersed in a social setting for some time in order to observe and listen with a view to gaining an appreciation of the culture of a social group (Bryman, 2016, p. 377).

#### **3.1.2.1.4 Case study**

It is defined as the study of an instance in action, such as a child, a class, a school, or a community. 'It provides a unique example of real people in real situations, enabling readers to understand ideas more clearly than simply by presenting them with abstract theories or principles' (Cohen et al., 2007, p. 253). The main benefit of a case study then is that the form is the one or few instances that usually allow the researcher to deal with the subtleties and intricacies of complex, social situations (Hoadjli, 2015).

#### **3.1.2.1.5 Grounded theory**

'In qualitative research, we may develop theory during the data collection process. This largely inductive method means that we are building theory from data or ground the theory in the data. Grounded theory adds flexibility and allows the data and theory to interact. This process also helps us remain open to the unexpected. We can change direction of study and even abandon the original research question in the middle of a project if we discover something new and exciting' (Neuman, 2014, p. 177).

#### **3.1.2.2 Quantitative designs**

With regard to quantitative designs, the focus should be on surveys and experiments, where answering research questions and hypotheses requires the investigation of the relationships between and among variables.

##### **3.1.2.2.1 Surveys**

According to Creswell (2014), 'a survey design provides a quantitative or numeric description of trends, attitudes, or opinions of a population by studying a sample of that population' (p. 155). In other words, a survey is a detailed study in which the researcher collects data on attitudes, opinions or impressions by polling a sub-group of a population. Then, from sample results, the researcher generalises or draws inferences to the population.

### **3.1.2.2.2 Experiments**

'In an experiment, investigators may also identify a sample and generalize to a population; however, the basic intent of an experimental design is to test the impact of a treatment (or an intervention) on an outcome, controlling for all other factors that might influence that outcome' (ibid, p. 156). Thus, to ensure that the change in the outcome is due to the treatment and not to other factors, investigators randomly assign individuals to two groups. One group receives a treatment, whereas the other one does not (ibid).

However, in most educational setting, random assignment of students is rarely possible. Consequently, researchers often use a 'quasi-experimental design'. Quasi-experiments are similar to true experiments, but they do not use random assignment to create comparisons (cited in Dörnyei, 2007).

### **3.1.2.3 A Mixed-methods design**

There have been several typologies for classifying and identifying types of mixed-methods strategies. For example, Creswell (2014) presents three basic mixed methods designs which are as follows:

- Convergent parallel mixed methods design;
- Explanatory sequential mixed methods design; and
- Exploratory sequential mixed methods design;

#### **3.1.2.3.1 Convergent parallel mixed-methods design**

It is probably the most familiar mixed-methods strategy which is based on the key assumption that both qualitative and quantitative data provide different types of information that should yield similar results. These information are often detailed views of participants qualitatively and scores on instruments quantitatively. Thus, in this approach, a researcher collects both quantitative and qualitative data, analyses them separately, and then compares the results to see if the findings confirm or disconfirm each other (ibid).

### **3.1.2.3.2 Explanatory sequential mixed-methods design**

Researchers with a strong quantitative background or from fields relatively new to qualitative approaches are mostly interested in this mixed-method design. This design includes two phases. In the first phase, the researcher collects quantitative data, analyses the results, and then uses the results to plan the second qualitative phase. The quantitative results typically determine the types of participants to be selected for the qualitative phase. They also determine the types of questions that will be asked of the participants. The purpose of this design is to have the qualitative data help explain in more detail the initial quantitative results. For example, a researcher might collect survey data in the first phase, analyse the data, and then use qualitative interviews to help explain the survey responses (ibid).

### **3.1.2.3.3 Exploratory sequential mixed-methods design**

Unlike the explanatory sequential approach, in the exploratory sequential approach, researchers start with a qualitative phase first followed by a quantitative phase. In other words, researchers begin by exploring with qualitative data and analysis. Then, they use the findings in a second quantitative phase. However, in both mixed-methods designs ; the second database builds on the results of the initial database. The intent of the exploratory sequential mixed-methods design is to develop better measurements with specific samples of populations. It also attempts to ensure that the data collected from a few individuals in qualitative phase can be generalized to a large sample of a population in quantitative phase (ibid).

There are also three advanced mixed-methods designs which are the embedded mixed-methods design, transformative mixed-methods design, and multiphase mixed-methods design. These strategies combine elements of the three basic mixed-methods designs discussed before (ibid).

### **3.1.3 Data collection methods**

Research methods are the tools used to collect data. They are 'a set of procedures that are intentional and planned for the purpose of collecting a certain type and number of data

sources that will be used to address a research question (Watkins & Gioia, 2015, p. 46). Social researchers can use four main methods of data collection, which are questionnaires, interviews, observation and documents. These tools help researchers to obtain:

- a clearer picture of things,
- an accurate measurement of things, and
- facts and evidence about the subject matter. (Denscombe, 2014, p. 163)

When selecting and using a specific research method, a researcher has to consider three main points. First, research methods are often associated with specific research strategies. Second, each method has its strengths and weaknesses ; so, a researcher has to use the method that serves best the purposes of the study. Third, a research method should not be considered as mutually exclusive. Thus, a researcher can use more than one method, so that the strengths in one method can compensate the weaknesses in another method (ibid).

### **3.1.3.1 Qualitative data collection methods**

In many qualitative studies, researchers spend a considerable time in the natural settings collecting multiple forms of data. For this reason, qualitative researchers use four basic types of data collection methods, which are observation, interviews, questionnaires, and focus group.

#### **3.1.3.1.1 Observation**

Observation is the data collection method that offers the researcher the opportunity to gather 'live' data from naturally occurring social situations. Hence, the researcher can directly observe what is occurring in the setting rather than relying on second hand accounts (Cohen et al., 2007)

#### **3.1.3.1.2 Interviews**

Maccoby and Maccoby (1954) define the interview as 'a face-to-face verbal interchange, in which one person, the interviewer, attempts to elicit information or expressions of opinion or belief from another person or persons' (cited in deMrrais & Lapan, 2004, p. 54). In other words, 'interviews enable participants to discuss their

interpretations of the world in which they live, and to express how they regard situations from their own point of view' (Cohen et al., 2007, p. 349).

### **3.1.3.1.3 Questionnaires**

'Questionnaires are any written instruments that present respondents with a series of questions or statements to which they are to react either by writing out their answers or selecting from among existing answers' (Brown, 2001, p. 06). The advantages of questionnaires are as follows:

- They can supply considerable amount of research data for relatively low cost in terms of time, money, and materials.
- They are simple and easy to administer.
- They provide standardised answers.
- They allow the speedy collection and analysis of data.

(Hoadjli, 2016, pp 44-45)

### **3.1.3.1.4 Focus group**

The focus group is the research method that 'is based on the collective experience of group brainstorming, that is, participants thinking together, inspiring and challenging each other, and reacting to the emerging issues and points' (Dörnyei, 2007, p. 144). Thus, this type of interaction can result in high-quality data.

### **3.1.3.2 Quantitative data collection methods**

In quantitative research studies, investigators can use two main data collection methods which are tests and structured questionnaires.

#### **3.1.3.2.1 Tests**

Cohen et al. (2007) regard tests as a powerful data collection method that is used to gather numerical data rather than verbal kinds. Before using testing for collecting research data, a researcher must consider the following issues:

- What are we testing (e.g. achievement, aptitude, attitude, personality,



intelligence, social adjustment etc.)?

- Are we dealing with parametric or non-parametric tests?
- Are they norm-referenced or criterion-referenced?
- Are they available commercially for researchers to use or will researchers have to develop home-produced tests?
- Do the test scores derive from a pretest and post-test in the experimental method?
- Are they group or individual tests?
- Do they involve self-reporting or are they administered tests? (ibid, p. 414)

### **3.1.3.2 Structured questionnaires**

According to Kothari (2004), structured questionnaires are those questionnaires which include definite, concrete and pre-determined questions. These questions may be closed or open questions. They are presented with exactly the same wording and in the same order to all respondents. Structured questionnaires may also have fixed alternative questions; as a result, the responses are limited to the stated alternatives. In this case, such type of questionnaire can be described as being highly structured. Hence, a highly structured questionnaire is one in which all questions and answers are specified and comments in the respondent's own words are held to the minimum.

### **3.1.3.3 A Mixed-methods data collection methods**

As it has been mentioned so far a mixed-methods research involves collecting both quantitative and qualitative data. This requires the use of both quantitative and qualitative data collection methods, such as tests, questionnaires, individual interviews, focus groups, and participants observations. Watkins and Gioia (2015) report that in a mixed-methods studies, researchers frequently choose the data collection methods during the planning and design phases. When deciding which data collection methods are needed, researchers must consider some key aspects which may influence their decisions. These aspects are, for example, the size of the sample, the scope of the study, the support and participation of units, resources, time, and previous research methods. Consequently, investigators can select the most appropriate data methods that will help them address the research questions.

### 3.1.4 Sampling

The quality of a research study depends on the appropriateness of methodology and instrumentation, as well as on the suitability of the sampling strategy which has been selected (cited in Cohen et al., 2007). Because it is difficult for researchers to gain information from the whole population due to several factors, such as expense, time and accessibility, they often need to collect data from a smaller group or subset of the total population. This smaller group or subset is called the sample.

Thus, 'sampling consists of selecting some part of a population to observe so that one may estimate something about the whole population (Thompson, 2012, p. 1). Likewise, a similar definition is that sampling is the process of selecting a few from a bigger group to become the basis for estimating or predicting the prevalence of an unknown piece of information, situation, or outcome regarding the bigger group (Kumar, 2011). In the same vein, Kothari (2004) also defines a sample design as, 'a definite plan for obtaining a sample from a given population. It refers to the technique or the procedure the researcher would adopt in selecting items for the sample. Sample design may as well lay down the number of items to be included in the sample i.e., the size of the sample' (pp. 55-56).

In social sciences, two methods of sampling are used which are probability and non-probability sampling. 'Probability sampling is based on the concept of random selection, whereas non-probability sampling is non-random sampling' (ibid, p. 58). That is, with probability sampling, every item in the population has an equal chance of being included in the sample, and its mathematical probability can be calculated. By contrast, with non-probability sampling, population elements are selected on the basis of their availability because they are volunteered, or because of the researcher personal judgment that they are representative. The consequence is that an unknown portion of the population is excluded.

The two methods of sampling can be categorised as follows:

Probability or random samples:

- Simple sample;
- Systematic sample;
- Stratified sample;
- Multi-stage sample; and
- Cluster sample.

Non-probability samples:

- Convenience sample;
- Purposive sample; and
- Quota sample.

## **3.2 Research methodology for this dissertation: Procedures**

### **3.2.1 Research approach**

The present study aims mainly at investigating the participants' attitudes towards the implementation of a Formal Oral Test, as well as the factors that contribute to these attitudes. It also seeks to explore the participants' perceptions of English language learning and assessment. Therefore, the research approach adopted to conduct this study is the qualitative approach because it is thought to be appropriate for such kinds of issues. In other words, it serves the nature of our research. Besides, the qualitative approach was selected because of its suitability to answer the research questions as it offers the appropriate data collection methods.

### **3.2.2 Research design / strategy**

A research strategy is a plan of action which is designed in order to answer the research questions and / or test the hypotheses. Thus, in the present research study, the strategy that is used is a case study. First, because there is a clear link between the purpose of the research and the chosen strategy. Also, we believe that it will be successful in achieving the aims of the present research as it will produce appropriate kinds of data. That is, it will produce findings that can answer the research questions. A case study is the strategy of inquiry that will enable us to study the group of participants in action and in a limited context.

### **3.2.3 Data collection methods**

As discussed before, when choosing the research method, the researcher has to take into account three main points, which are the association between research methods and specific research strategies, the purposes of the study, and the need for more than one method. Therefore, the data collection methods selected for the present research include a questionnaire and an interview.

### **3.2.3.1 Questionnaires**

#### **3.2.3.1.1 Aim**

With regard to the questionnaire, a preliminary questionnaire was first employed in order to explore the pupils' perceptions of English language learning and assessment. Then, a final questionnaire was used after the pupils took a speaking test. The main purpose of this final questionnaire was to investigate the pupils' attitudes towards the implementation of a Formal Oral Test. Both questionnaires were designed in English language. However, it is important to mention that they were translated into Arabic language, the native language of the participants, before they were administered to the pupils. The reason behind this was to avoid any misunderstanding caused by the language factor, and to make the respondents generate their ideas to the maximum.

#### **3.2.3.1.2 Structure and content**

The preliminary questionnaire consisted of two parts. Section One included 15 questions. The questions were a combination of both closed-ended and open-ended questions. The first three questions were intended to explore the pupils' perceptions of the most important language skill. It also aimed at examining pupils' awareness of the importance of all aspects of English language learning. The questions from four to 10 were mainly devoted to explore the participants' perceptions of English language assessment. Then, questions 11 and 12 sought to gather information about the pupils' speaking and writing abilities. Finally, the rationale behind the three last questions was to investigate the participants' attitudes towards the implementation of a Formal Oral Test before they experienced a speaking test. Also, question 15 attempted to identify the pupils' favourite speaking assessment format.

On the other hand, Section One of the final questionnaire included 12 questions. They were also a combination of both closed-ended and open-ended questions. The main purpose of the questions from one to eight was to collect the participants' opinions of the speaking test which they took before they received the questionnaire. They attempted to access the thoughts and feelings that the respondents experienced during the speaking test. Besides, question eight sought to identify the respondents' attitudes towards group

assessment, and whether or not it provided them with an effective and comfortable context. Then, the questions nine and 10 aimed at examining the participants' opinions of the most important language skill, as well as their awareness of the significance of all English language aspects. These two questions were meant to compare the participants' responses to the ones they provided in the preliminary questionnaire. Finally, the last two questions were devoted to investigate the pupils' attitudes towards the implementation of a Formal Oral Test.

Section Two of both questionnaires was developed in order to enable the pupils to express their opinions on the the contents and layout of the questionnaires. It consisted of six questions. Most of them were yes-no questions. This part stood as an opinionnaire only in the piloting stage. That is, it was omitted in the final version.

### **3.2.3.1.3 Piloting and validation**

To increase the reliability, validity and practicability of the preliminary and final questionnaires, everything about them should be piloted. That is, piloting is a crucial step to check the comprehensibility of the questions, to eliminate any ambiguities or difficulties in wording, and to gain feedback on the appearance of the questionnaires. Also, we aimed at checking the time taken to complete each questionnaire. The piloting process was carried out two days before the final version of each questionnaire was administered to the participants.

In the piloting stage, the preliminary questionnaires were handed to five pupils who were chosen randomly from our population. The pupils took about 15 to 20 minutes to answer all the questions. They all agreed on the clarity of the questions, and the attractiveness of the layout.

The same steps were also followed when piloting the final questionnaire, and we got the same feedback. It is important to note that both questionnaires were administered to our supervisors who said that no changes were required.

### **3.2.3.2 Interview**

#### **3.2.3.2.1 Aim**

The interview was employed because this data collection method allows direct contact with the sample. It is flexible. It produces in-depth details of the topic. The general aim of the interview was to elicit more information about the participants' attitudes towards the implementation of a Formal Oral Test. We believed that we could collect more data because the participants could talk freely about their experience with the speaking test. The interviews aimed at exploring the participants' feelings when they were taking the test, as well as the difficulties that they faced when they were carrying out its different tasks. In addition, the interviewees were also expected to provide their opinions of group assessment format.

#### **3.2.3.2.2 Structure and content**

The researchers used the final questionnaire as a guide to design the interview which included 11 questions. We decided to use semi-structured interviews. The first two questions were meant to set the tone, and create a comfortable atmosphere. The reason behind this was to encourage the interviewees to open up. Then, the next seven questions were the main ones. They aimed at collecting more useful information about the topic of the present study. The last two questions were the final closing questions which were: 'Do you want to add anything?', and 'Do you have any questions that you want to ask?'. These questions permitted the interviewees to add any comments, or say anything that they might forget to mention during the interviews.

### **3.2.4 Data collection procedures**

#### **3.2.4.1 Data collection procedures for the questionnaires**

The preliminary questionnaires were handed to the 36 participants in a session of English language. All the participants responded to the questionnaires in the same session. Then, after taking a speaking test, the participants were given the final questionnaires. We followed exactly the same procedures when collecting data using the final questionnaire.

### 3.2.4.2 Data collection procedures for the interview

Nine pupils were randomly chosen from the sample to be interviewed. The participants were informed about the purpose of the interview in advance, and also at the beginning of each interview. The interviews were conducted in the target language. However, it is worth mentioning that the interviewer sometimes translated some questions into the participants' native language, so that the interviewed pupils could better understand the questions. Even the interviewees were allowed to speak in Arabic language whenever they felt unable to express their ideas in English language.

The researcher raised 11 questions. The two first questions were asked in order to create a familiar and comfortable atmosphere. Then, the participants were asked the seven main questions. The interviewer sometimes used probes in order to increase the richness and depth of the responses. At the end, the researcher asked the two final closing questions. All the interviews were audio recorded. The time devoted to each one was between four to seven minutes. We could not interview the pupils for a longer time because they had other classes to attend. After finishing the interviews, the latter were transcribed into written versions.

**Table 3.1:** The Order of using data collection methods

<b>Stage One:</b> Exploring pupils' perceptions of English language learning and assessment <b>Data collection method:</b> Preliminary questionnaire
<b>Stage Two:</b> English speaking test
<b>Stage Three:</b> Investigating pupils' attitudes towards the implementation of a Formal Oral Test. <b>Data collection methods:</b> Final questionnaire and interview

### 3.2.5 Population and sampling technique

The population of this study was fourth year pupils at Cid Nouredine Middle School in Tolga, Biskra (Algeria). It consisted of 111 pupils. One class, which included 36 pupils, was chosen to participate in this investigation. Thus, the sampling technique which was selected for the present research is a purposive sample. The characteristics of this method of sampling are as follows:

- It is used by some arbitrary method because it is known to be representative

of the total population; or it is well known that it will produce well matched groups.

- Its main idea is to pick out the sample in relation to some criteria, which are considered important for particular studies.
- This technique is appropriate when the study places special emphasis upon the control of specific variables. (Hoadjli, 2016, p. 52)

### **3.2.6 Data analysis procedures**

#### **3.2.6.1 Data analysis procedures for the questionnaires**

The data analysis procedures for both questionnaires were the same. The collected data were analysed following the various steps of qualitative data analysis. With regard to closed-ended questions, all the percentages were calculated manually, and tables were created in Microsoft® Word.

#### **3.2.6.1 Data analysis procedures for the interview**

The data obtained from the interview were also analysed following the different steps of content-based approach. As it is mentioned before, all the interviews were audio-recorded. So, the audio recordings were first transcribed. That is, the full script of each interview was produced. Then, the scripts were printed out. Next, the necessary parts of scripts were marked using coloured highlighters. After that, the main categories were identified, as well as the relationships between them. Finally, the results were presented as a descriptive and interpretive account of the data.



### 3.3 The Results

#### 3.3.1 Preliminary questionnaire

**Item1:** *Why do you study English?*

**Table 3.2:** Pupils' reasons for studying English language

Option	Number	Percentage
To satisfy school requirements.	05	13.88%
To be able to pass the Brevet Exam.	07	19.44%
To be able to communicate with foreigners.	13	36.11%
Other reasons	11	30.55%

This question is intended to seek information about what motivates pupils to study English language. An analysis of the results above indicates that 36.11% of the pupils study English language basically in order to be able to communicate with people from other countries, and this is , of course, the main purpose of learning any foreign language. While 30.55% of the pupils claimed that they studied English language for other reasons, the rest said that they studied English language because they had to, or just in order to be able to pass the Brevet Examination and follow their studies at high school.

**Item 2:** *In your view, what aspects of English language is the most important one?*

**Table 3.3:** Pupils' awareness of the importance of all aspects of English language

Option	Number	Percentage
Vocabulary	10	27.77%
Grammar	02	05.55%
Pronunciation	08	22.22%
All the above	16	44.44%

This item aims at determining whether pupils are aware of the importance of each aspect of English language or not. The largest number of the pupils, which is 16 pupils ( 44,44%) know that learning a foreign language means giving attention to all aspects of language learning. On the other hand, the rest of the pupils stated that the focus should

only be on one of the components of English language. While, 10 pupils (27,77%) believed that EFL learners should seek to memorise a great number of words and expressions, eight pupils (22,22%) said that pronunciation was the most important aspect of English language, and only two pupils (05,55 %) thought that mastering the grammatical rules should be a priority.

**Item 3:** *Please rank the following language skills in order of importance from 1 to 4 where 1 is most important to you and 4 is least important to you.*

**Table 3.4:** Pupils' perceptions of language skills

Option	Number				Percentage			
	1	2	3	4	1	2	3	4
Listening	12	09	08	07	33.33%	25%	22.22%	19.45%
Speaking	17	05	05	09	47.22%	13.89%	13.89%	25%
Reading	06	14	11	05	16.66%	38.89%	30.56%	13.89%
Writing	01	08	12	15	2.78%	22.22%	33.33%	41.67%

The aim of this question is to determine pupils' perceptions of language skills. The results show that a large number of the respondents (47.22%) believed that speaking was the most important language skill. In addition, 33.33% of the respondents thought that listening was more important than reading and writing. However, 38.89% gave more importance to reading than listening and writing. Moreover, 41.67% of the respondents considered writing as the least important skill.

**Item 4:** *Do achievement tests cover the content of what you have learnt?*

**Table 3.5:** Correspondence between test-contents and syllabus

Option	Number	Percentage
To a great extent	25	69.44%
To some extent	10	27.77%
To a very little extent	00	00.00%
No correspondence between the two	01	02.78%

This question sheds light on the relationship between test-contents and the syllabus.

The results show that while 25 pupils (69,44%) out of the 36 respondents thought that achievement tests covered a great extent of the lessons they have been taught in the classroom, 10 pupils (27,77 %) saw that the test-contents and the syllabus correspond only to some extent. However, only one pupil (02,78%) believed that there was no correspondence between them.

**Item 5:** *How are the testing methods used in these tests?*

**Table 3.6:** Testing methods variety

Option	Number	Percentage
The same in every test.	22	61.11%
Vary slightly from one test to another	14	38.89%
Vary completely from one test to another	00	00.00%

As for this item, the majority of the respondents (61.11%) asserted that the testing methods used to assess their achievement and progress were always the same. However, 38.89% of the pupils stated that the testing methods varied slightly from one test to another.

**Item 6:** *What do you think the currently used tests assess?*

**Table 3.7:** The tested ability

Option	Number	Percentage
Your degree of comprehension	14	38.89%
The amount of vocabulary you know.	03	08.33%
Your mastery of grammatical structures.	06	16.67%
Your ability to communicate.	01	02.78%
All the above	12	33.33%

This item aims at determining what the currently used tests are intended to measure. While 02.78% of the respondents reported that the achievement tests assessed their ability to communicate, 38.89% of the respondents said that these tests were used to assess their

degree of comprehension. 33.33% of the respondents thought that the currently used tests tested all aspects of language, namely vocabulary, grammar and pronunciation, 16.67% believed that tests assessed their mastery of grammatical rules, and 08.33% saw that tests measured the amount of vocabulary pupils knew.

**Item 7:** *How is the distribution of test-items among the test parts?*

**Table 3.8:** Test-items distribution

Option	Number	Percentage
Equal	06	16.67%
Not equal	20	55.55%
It depends from one test to another.	10	27.78%

The results show that 55.55% of the pupils noted that the distribution of test-items among the test parts was not equal. The currently used tests include two section. Section one is divided into two parts which are 'Reading Comprehension' and 'Mastery of Language', and Section Two is 'Written Expression'. The number of test-items among these parts is always different. However, 27.78% of the respondents said that the distribution of test-items varied from one test to another ; whereas, 16.67% thought that the distribution of items was equal.

**Item 8:** *Which language skills are almost neglected in the currently used tests?*

**Table 3.9:** Neglected language skills in achievement tests

Option	Number	Percentage
Listening	20	30.77%
Speaking	29	44.62%
Reading	10	15.38%
Writing	06	09.23%

This question is intended to discover whether the pupils are aware of the language skills which are almost neglected in achievement tests or not. The analysis of the results above reveals that 44.62% of the respondents asserted that speaking was not tested, and 30.77% of the respondents reported that listening was completely neglected. However,

15.38% of the pupils claimed that reading was the skill which was given no importance ; whereas, 09.23% thought that teachers did not test the writing skill.

**Item 9:** *Do you think the contents of these tests correspond to your interests?*

**Table 3.10:** Correspondence between test content and pupils' interest

Option	Number	Percentage
Yes.	10	27.78%
No.	15	41.67%
Somehow.	11	30.55%

This item sheds light on pupils' attitudes towards the contents of achievement tests. 41.67% of the pupils questioned saw that the contents of the currently used tests did not correspond to their interests, and 30.55% reported that test contents were somehow in accordance with their interests. The pupils who chose the second or the third option provided mainly two reasons. They asserted that achievement tests did not assess pupils' speaking abilities. They assessed only pupils' degree of comprehension, their mastery of grammatical rules, and some written sub-skills. These pupils also wrote that even the topics that tests dealt with were sometimes not interesting. However, 27.78% of the respondents believed that there was a match between test contents and their interests. They claimed that achievement tests covered various topics and that they are based on the lessons presented in the classroom. Thus, it can be deduced that teachers focus basically on teaching grammar and vocabulary rather than developing pupils' language skills.

**Item 10:** *Do your marks in tests usually reflect your progress and achievement in English language?*

**Table 3.11:** Correlation between test scores and pupils' level

Option	Number	Percentage
Yes.	11	30.56%
No.	12	33.33%
To some extent.	13	36.11%

As for this item, what is noticed is that the responses of the pupils are completely

different. 30.56% of the pupils believed that their grades in achievement tests really reflected their progress and achievement in English language, whereas, 33.33% of the pupils saw that the grades they obtained in achievement tests did not completely reflect their level. However, 36.11% of the respondents thought that there is some kind of correlation between their test scores and their progress.

**Item 11:** *To what extent you are able to express yourself through writing?*

**Table 3.12:** Pupils' writing abilities

Option	Number	Percentage
To a great extent	07	19.44%
To some extent	18	50%
To a very little extent	11	30.56%

This question aims at discussing pupils' abilities to express themselves through writing. The analysis of the results above shows that half of the pupils claimed that they had good writing skills, but they have some problems, such as lack of vocabulary and spelling difficulties. As a result, they sometimes found it difficult to express all their ideas through writing. 30.56% of the respondents reported that they had poor writing skills ; so, they struggled when they attempted to write paragraphs. However, 19%.44 of the respondents asserted that they were able to express themselves well through writing. They wrote that they did not face any difficulties because they had rich vocabulary and they master different grammatical structures.

**Item 12:** *To what extent are you able to express yourself through speaking?*

**Table 3.13:** Pupils' speaking abilities

Option	Number	Percentage
To a great extent	08	22.22%
To some extent	20	55.56%
To a very little extent	08	22.22%

This item aims at discussing pupils' abilities to express themselves through speaking. The analysis of the results above indicates that 55.56% of the pupils said that they were

able to speak English language in an acceptable way. However, they reported that they had some pronunciation problems, and they couldn't sometimes find the right words to express their ideas.

Concerning the rest of the pupils, half of them stated that they could speak English Language fluently because they started learning English language at an early age. Besides, they spoke English language whenever they had an opportunity to do so. The other half reported that because of many reasons, they faced great challenges when they attempted to communicate with other people. They wrote that they lacked the necessary vocabulary to express their ideas. Also, they did not know how to pronounce words correctly. Moreover, they had poor communication skills as they did not get enough opportunities to speak English language whether inside or outside the classroom. Thus, it can be said that teachers do not make enough efforts to develop pupils' speaking abilities; that is, they do not pay enough attention to teaching the speaking skills.

**Item 13:** *In your opinion, should listening and speaking skills be given more importance in English language tests?*

**Table 3.14:** Pupils' opinions about the importance of listening and speaking assessment

Option	Number	Percentage
Yes.	26	72.22%
No.	10	27.78%

The aim of this item is to identify pupils' attitudes towards the idea of giving more importance to listening and speaking skills in achievement tests in order to enhance English language assessment. The analysis of the results above shows that the majority of the pupils (72.22%) agreed that listening and speaking should be given more attention as they were two basic language skills. They asserted that they had to develop these skills because the main reason for learning a foreign language was to be able to communicate with people from other countries. They added that listening and speaking were the most commonly used skills. Moreover, for some pupils, it is easier to express themselves through speaking than it is through writing. Therefore, tests should assess the four skills in order to evaluate learners' achievement and progress.

On the other hand, 27.78% of the respondents did not like the idea of giving more importance to listening and speaking skills in achievement tests. They claimed that the

currently used tests were enough to assess pupils' knowledge and skills. In addition, they thought that the most important thing was to learn the grammatical rules and to be able to apply them correctly. They reported that they did not need to improve their listening and speaking skills in order to succeed in their studies.

**Item 14:** *Do you think that the implementation of a Formal Oral Test will be a good idea to improve English language assessment?*

**Table 3.15:** Pupils' attitudes towards the implementation of a Formal Oral Test

Option	Number	Percentage
Yes.	18	50%
No.	18	50%

This item aims at discussing pupils' attitudes towards the implementation of a Formal Oral Test. The results reveal that while half of the pupils believed that the implementation of a Formal Oral Test would be a good idea to improve English language assessment, the other half claimed that the incorporation of a Formal Oral Test was not a necessity.

On the one hand, those pupils who thought that the implementation of a Formal Oral Test was a good idea gave several reasons. First, they believed that it would contribute to the development of pupils' communication skills as they were children. Second, it would build pupils' self-confidence. Finally, because there were some pupils who could express themselves better through speaking than they could through writing, the implementation of a Formal Oral Test would enhance English language assessment.

On the other hand, the respondents who thought that the incorporation of a Formal Oral test was not a good idea provided different reasons. First, they said that they did not speak English well ; so, a speaking test would not be useful for them. Second, they said that pupils must focus on reading and writing skills, and achievement tests should be written not oral. Finally, they believed that a Formal Oral Test would make things harder for pupils as it would increase the demands of success.



**Item 15:** *How would you like your speaking performance to be assessed?*

**Table 3.16:** Pupils' favourite speaking assessment format

Option	Number	Percentage
Individually	16	44.44%
In pairs	06	16.67%
In groups	14	38.89%

The aim of this question is to identify the speaking assessment format that the pupils would prefer. 44.44% of the respondents said that they would prefer to be assessed individually rather than being assessed in pairs or groups. Some of them claimed that they were shy and that they would be unable to communicate their ideas if they were assessed with other pupils. Others said that assessing them in groups would cause stress and anxiety, and that they would feel embarrassed if they made any mistakes. However, 38.89% of the pupils questioned reported that group format would provide an effective and a comfortable context for speaking assessment, while 16.67% of the pupils thought that paired format would be better. The pupils who preferred group or paired assessment said that they would be able to interact with each other and share their experiences and ideas, and this would, in turn, contribute to their overall assessment. Also, they believed that they would learn new information and gain knowledge.

### 3.3.2 Final questionnaire

**Item 1:** *How was the speaking test?*

**Table 3.17:** Pupils' opinions about the English speaking test

Option	Number	Percentage
Very easy	11	30.56%
Easy	12	33.33%
Neutral	07	19.44%
Difficult	02	05.56%
Very difficult	04	11.11%

This item seeks to identify pupils' opinions about the English speaking test. The results show that the majority of the respondents claimed that the test was easy as 30.56% of them chose the first option and 33.33% chose the second one. 19.44% of the respondents saw that the speaking test was neither easy nor difficult. However, 05.56% of the pupils claimed that the test was difficult, and 11.11% thought that it was very difficult.

**Item 2:** *Were the questions/tasks the ones you had never encountered or expected?*

**Table 3.18:** Pupils' experiences of the questions of the speaking test

Option	Number	Percentage
Yes.	12	33.33%
No.	15	41.67%
Somehow.	09	25%

This question aims at seeking information about the questions that the pupils were asked to answer and the tasks that they were asked to perform in the speaking test. 41.67% of the pupils said that they had never been asked such questions before, and that they did not expect them. However, 25% of the respondents reported that they did not expect some of the questions, but the others were easy and familiar ones. On the other hand, 33.33% asserted that the questions and the tasks were common, and that they had already experienced them.

**Item 3:** *Were the questions/tasks in the speaking test explicit and clear?*

**Table 3.19:** Pupils' opinions about questions and tasks of the speaking test

Option	Number	Percentage
Yes.	30	83.34%
No.	03	08.33%
Somehow.	03	08.33%

This question is intended to determine pupils' opinions about the questions and the tasks which were in the speaking test. As it is noticed from the table above the majority of the respondents (83.34%) reported that the questions and tasks were easy and clear. However, 08.33% of the respondents claimed that some of the questions were not clear, and 08.33% thought that the questions were ambiguous. We believe that the pupils who reported that

the questions were not clear or somehow ambiguous are the pupils who do not master the English language.

**Item 4:** *How did you feel during the speaking test?*

Most of the pupils we questioned wrote that they felt stressed and anxious because they were taking an English speaking test for the first time in their lives. However, many of them stated that they started to feel comfortable and more confident after a few minutes. Others said that they felt happy and excited, especially when they started to interact with other pupils. In general, most if not all of the the pupils reported that they liked this new experience.

**Item 5:** *Did the assessor encourage and motivate you during the speaking test?*

**Table 3.20:** Pupils' opinions of the Assessor

Option	Number	Percentage
Yes.	23	63.89%
No.	06	16.67%
Somehow.	07	19.44%

The aim of this question is to determine pupils' opinions of the assessor. The table above shows that the majority of the pupils (63.89%) asserted that the assessor motivated and encouraged them during the test. Thus, it can be deduced that the examiner created a comfortable atmosphere for the speaking assessment by providing the pupils with sufficient encouragement. However, 19.44% of the respondents reported that they did not get enough encouragement, and 16.67% of the respondents claimed that they did not receive any encouragement from the assessor. We believe that maybe those pupils were confident of themselves, and they did not need to be motivated.

**Item 6:** *What were the difficulties that you faced during the test?*

The majority of the pupils wrote that the main problem that they faced during the speaking tests was how to say some words in English language. They added that because of their lack or shortage of vocabulary, they were unable to form correct sentences and express their ideas. They also asserted that they had some difficulties with pronunciation. Others said that they had poor communication skills; as a result, they could not interact

effectively with the other pupils. In addition, some pupils wrote that they faced great challenges when they attempted to understand or answer the questions.

**Item 7:** *What did the speaking test allow you to do?*

Many pupils said that the speaking test was the most important test to encourage them to use English language. It allowed them to communicate with each other, and to share their ideas, opinions and experiences. They added that they believed that speaking tests were necessary to discover their levels in English language. Other pupils wrote that the test motivated them to study English language and improve their speaking abilities because they realised that speaking is the most important skill to be developed.

**Item 8:** *According to you, does group format provide an effective and a comfortable context for speaking assessment?*

**Table 3.21:** Pupils' attitudes towards group assessment

Option	Number	Percentage
Yes.	28	77.78%
No.	04	11.11%
Somehow.	04	11.11%

This item aims at identifying pupils' attitudes towards group assessment after they took an English speaking tests and experienced group format. When comparing the results above with the ones obtained before, we notice that the percentages have changed.

The results above show that the majority of the respondents saw that group format provide an effective and a comfortable context for speaking assessment. They reported that group assessment could reduce stress levels and test pressure that pupils may feel. In addition, they claimed that assessing pupils in groups could motivate them to speak as they would like to interact with each other and share ideas and information. Moreover, they believed that group assessment would help pupils build their self-esteem and self-confidence.

Some respondents (11.11%) thought that it was better to reduce the number of the pupils and use paired format. However, 11.11% of the respondents did not change their opinions, and they still believed that it was better to assess each pupil alone because group assessment could cause stress and anxiety. They added that some pupils may also feel

embarrassed if they made any mistakes.

**Item 9:** *Please rank the following language skills in order of importance from 1 to 4 where 1 is most important to you and 4 is least important to you*

**Table 3.22:** Pupils' opinions about language skills

Option	Number				Percentage			
	1	2	3	4	1	2	3	4
Listening	14	08	06	08	38.89%	22.22%	16.67%	22.22%
Speaking	16	09	07	04	44.45%	25%	19.44%	11.11%
Reading	06	10	16	04	16.67%	27.78%	44.44%	11.11%
Writing	00	09	07	20	00.00%	25%	19.44%	55.56%

This item aims at identifying pupils' opinions about the importance of each language skill after they took an English Speaking Test. The analysis of the results above show that the respondents did not change their opinions. The majority of the respondents still believe that speaking and listening are the basic language skills because 44.45% of the respondents chose speaking as the most important skill and 38.89% chose listening. In addition, 44.44% of the respondents claimed that reading is less important than speaking and listening, but more important than writing. Moreover, 55.56% of the respondents considered writing as the least important skill.

**Item 10:** *In your view, what aspects of English language is the most important one?*

**Table 3.23:** Pupils' realisation of the importance of all aspects of English language

Option	Number	Percentage
Vocabulary	07	19.45%
Grammar	03	8.33%
Pronunciation	04	11.11%
All the above	22	61.11%

The aim of this item is to determine whether the pupils have become aware of the importance of each aspect of English language or not. We can notice that the percentages

have changed after the pupils took the speaking test. The largest number of them have realised the importance of each aspect of English language as the results show that 61.11% of the respondents chose the last option. However, 19.45% of the respondents still thought that learning English language meant learning a large number of words. Also, 11.11% chose pronunciation as the most important aspect, and 8.33% claimed that EFL learners should focus on mastering the grammatical rules.

**Item 11:** *Now, do you think that your marks in currently used tests usually reflect your progress and achievement in English language?*

**Table 3.24:** Correlation between test scores and pupils' level

Option	Number	Percentage
Yes.	06	16.66%
No.	19	52.78%
To some extent.	11	30.56%

This item was intended to determine whether the pupils still could not decide if their test scores reflected their achievement and progress in English language. What is noticed is that 52.78% of the respondents reported that they believed that achievement tests did not really assess their knowledge and skills. In addition, 30.56% of the respondents saw that their tests grades somehow reflected their level in English language. They wrote that they thought so mainly because achievement tests did not test all language skills. That is, currently used tests did not measure their listening and speaking skills, despite the fact that they are the most important language skills. Also, some respondents reported that there were some pupils who had good speaking abilities, but written tests did not allow those pupils to show them.

On the other hand, 16.66% of the respondents still believed that there was a correlation between their test scores and their progress since achievement test covered the lessons presented in the classroom. Therefore, we can say that teachers teach to the test, and they do not make enough efforts to develop pupils' speaking skills.

**Item 12:** *Do you think that speaking test should be administered in a formal context for grades?*

**Table 3.25:** Pupils' attitudes towards the implementation of a Formal Oral Test

Option	Number	Percentage
Yes.	22	61.11%
No.	14	38.89%

This item aims at discussing pupils' attitudes towards the implementation of a Formal Oral Test after they took a speaking test. The results reveal that some pupils have changed their attitudes, and that they believe now that taking a Formal Oral Test is a necessary step to improve English language assessment. The table above shows that 61.11% of the respondents agreed with the idea that a speaking test should be administered in a formal context for grades. First, because they believed that it would develop pupils' self-confidence and self-esteem. Second, it would improve their speaking skills. Third, it would improve pupils' tests scores since some of them had good speaking abilities. Finally, the implementation of a formal Oral Test would enhance English language assessment.

On the other hand, 38.89% of the respondents claimed that the incorporation of a Formal Oral test was not a good idea. They saw that speaking tests were not important and that written tests were better. They believed that pupils must focus only on reading and writing skills in order to succeed. In addition, they claimed that during a speaking test pupils were under great pressure and this hindered their performance. Finally, some pupils said that because many pupils did not speak English language well, they would get bad grades. Thus, the implementation of a speaking test would not be useful for them.

### 3.3.3 Interview

This part of the research presents findings from the interview.

The first two questions, as it was mentioned earlier, were intended in order to create a familiar and comfortable atmosphere for the interviewees.

**Question 3:** *There are four language skills which are listening, speaking, reading and writing. In your opinion, which language skill is the most important one?*

Concerning this question, there was a general agreement among the interviewees

who asserted that speaking is the most essential skill as they believed that it is the most used one besides, of course, listening. Thus, **pupil 2** said, '*...it is important to make your listener understand you*'. In addition, **pupil 3** reported: '*Because, unlike writing, when I speak I don't need rules....When I write, I must write correct English without making mistakes. There is no problem if you make mistakes when you speak. Because we can explain our idea.*'

Therefore, it can be said that pupils, who have difficulties in communicating their ideas through writing, may find it easier to communicate their ideas through speaking because they do not have to pay a lot of attention to the grammatical rules or the punctuation marks.

**Question 4:** *How can you describe the speaking test which you took a couple of days ago?*

As for this question, the interviewees were asked to express their opinions about the speaking test. All of them agreed on that it was easy. They also claimed that the questions and tasks were clear and easy. They added that the speaking test was really an interesting experience.

**Question 5:** *How did you feel during the speaking test?*

This question was used to seek information about the psychological situation of the pupils during the tests. The majority of the interviewed pupils said that they were stressed and nervous at least at the beginning of the interview. Others said that they were comfortable and excited, and that they enjoyed it a lot.

**Question 6:** *What were the difficulties that you faced during the test?*

This question was raised by the researchers in order to identify the difficulties that the pupils faced during the speaking test. Some of the interviewees claimed that they did not face any difficulties at all. Others said that they faced two main problems which are lack or shortage of vocabulary and pronunciation problems. These are samples of some interviewees' responses:

**Pupil 8:** '*I couldn't say all what I wanted to say. I couldn't sometimes say some words in English.*'

**Pupil 7:** '*I didn't know how to say some words in English. I could not express myself in English.*'

**Pupil 6:** '*The pronunciation of some words.*'

In addition, **pupil 6** reported, '*Well, my partners did not communicate with me a lot. I*



*was the only one who was speaking. I asked them but...'* This means that some pupils have poor communication skills, and that they are unable to interact with other people in English language.

**Question 7:** *What did the speaking test allow you to do?*

Regarding this question, the interviewees provided different answers. Some of the interviewed pupils said that the speaking test allowed them to interact with other people in English language, and to express themselves orally. Thus, they have realised that speaking is the most important skill that EFL learners have to develop. Also, **pupil 6** reported, *'I was able to use the words that I know them, and to build my self-confidence.'* which means that speaking tests can even play an important role in the development of children's personality. Another interviewed pupil claimed that he learnt new information. Moreover, **pupil 3** said, *'I discovered that I need to learn more words.'* Therefore, it can be said the speaking test was really a good experience that enabled pupils to discover their strengths and weaknesses, and to learn new things whether about English language or life.

**Question 8:** *In your opinion, does group format provide an effective and a comfortable context for speaking assessment? Explain.*

The aim of this question was to determine the interviewees' attitudes towards group assessment. All of the pupils questioned said that group format can provide an effective and a comfortable context for speaking assessment. For example, **pupil 7** reported, *'I feel comfortable when I take the test with other pupils, not alone.'* and **pupil 1** said, *'...It helped me to take the speaking test.'* In addition, most of the interviewees asserted that group format allowed them to interact and communicate with each other, and share ideas and information. For example, **pupil 3** said, *'We can interact with each other, communicate with each other, and share some ideas.'*

**Question 9:** *Do you think that speaking test should be administered in a formal context for grades? And why?*

As for this question, the majority of the interviewees agreed on with the idea that the speaking test should be administered in a formal context for grades. They believed that it would contribute to the improvement of pupils' speaking skills. These are some of the interviewees' replies:

**Pupil 7:** *'We should learn to speak English fluently right from the beginning as we*

*are young...when we are children.'*

**Pupil 8:** *'It will help a lot of pupils to enhance their speaking abilities.'*

**Pupil 1:** *'It helped me to speak English.'*

Some of the questioned pupils also said that currently used tests do not assess all language skills. **Pupil 2** reported, *'Written exams do not assess all our abilities / skills. The speaking test allowed us to say things that we can not write. We can express ourselves in speaking tests better than in written ones. We can't write everything.'* Moreover, other pupils believed that speaking tests may help them get extra-points and increase their scores in English language.

### **3.4 Discussion of the Results**

To recapitulate, the main purpose of the present study is to investigate pupils' attitudes towards the implementation of a Formal Oral Test, and to identify the factors that contribute to their attitudes. Also, the study sought to explore pupils' perceptions of English language learning and assessment. In addition, we intended, through this study, to determine how pupils perceive their experience of group speaking assessment.

The findings drawn from the analysis of the data obtained using two data collection methods were positive in many respects. First of all, the preliminary questionnaire revealed that many pupils are aware of the importance of all aspects of English language, namely vocabulary, grammar and pronunciation. In other words, pupils know that they have to consider all the components of the English language, and that the focus should not only be on one or two of them. Surprisingly, the results obtained show that a large number of pupils consider speaking as the most important skill that they have to master. Therefore, many of the pupils reported that what motivated them to learn the English language was the desire to be able to communicate with foreigners. They study English language not only to satisfy school requirements, and to be able to pass the BEM Examination and go to high school. Many of the pupils also want to improve their communication skills because they believe that they are so important in their daily lives.

With regard to English language assessment, the percentages show that the currently used tests mainly assess pupils' degree of comprehension, as well as their mastery of grammatical rules. Additionally, the language skills which are almost neglected in achievement tests are speaking and listening. It is worth mentioning that only two points are devoted to pronunciation. Hence, the majority of pupils are not satisfied with their test

grades. They claimed that written tests did not really assess their progress and achievement in English language. Taking into account their age, many pupils believe that they have good speaking abilities, but written tests do not allow these pupils to demonstrate them.

Furthermore, the results show that the majority of pupils thought that tests covered a great extent of the lessons they dealt with in the classroom. Consequently, it can be concluded that teachers focus mainly on teaching grammar and vocabulary rather than developing pupils' language skills. It can be argued that teachers give no importance to teaching the speaking skill. They do not provide learners with enough opportunities to enhance their communication skills.

Moreover, this study revealed that the majority of pupils agreed that listening and speaking should be given more importance in English language tests. They were convinced that these two skills were the most basic language skills. Accordingly, half of the respondents to the preliminary questionnaire liked the idea of implementing a Formal Oral Test for several reasons. First, they claimed that it would contribute to the development of pupils' communication skills. Second, it would build pupils' self-confidence and self-esteem. Finally, the implementation of a Formal Oral Test will improve English language assessment and testing.

With reference to the preliminary questionnaire, pupils initially proposed that raters should assess them individually or in groups. Those who preferred individual assessment explained that they were shy, and they would be unable to speak in English language in front of other pupils. On the other hand, those who chose group assessment stated that it would be better than individual and paired assessment.

Concerning the speaking test that the participants took in the current study, the results obtained from the final questionnaire and interview show that more than half of the pupils agreed on that it was easy. They reported that the questions and tasks were clear and easy. All in all, it was a new and interesting experience for the pupils. Besides, many pupils said that the speaking test was the most important test to encourage them to use English language. It allowed them to communicate with each other, and to share their experiences, ideas, and opinions. As a result, the pupils now believe that oral tests are necessary to assess and improve their level of English language as it will enable them to discover their strengths and weaknesses. Moreover, speaking tests can play an important role in the development of children's personalities.

In addition, the information collected through this research indicate that most of the

pupils were a little bit stressed during the speaking test. Because they were taking a speaking test for the first time in their lives, the pupils were anxious, especially at the beginning of the test. However, many of them said that they started to feel comfortable and more confident after a few minutes. Some of the pupils claimed that they were happy and excited as they were speaking to each other in English language. Generally, it seemed that the pupils enjoyed this experience.

Furthermore, this study revealed the main problems that middle school pupils face when speaking. First and foremost, it can be said that the shortage or lack of vocabulary is a major problem among learners. This leads, of course, to their inability to form correct sentences; and consequently, they can not express their ideas. Besides, they have some pronunciation problems. In other words, they are unable to pronounce some English sounds, and they have problems with word stress and intonation. Thus, some pupils have poor communication skills, and they can not effectively interact with other people.

In terms of their perceptions of the experience of the group assessment, the majority of the pupils indicated that group format provide a comfortable context for speaking assessment. They think that group assessment can reduce stress levels and test pressure. Additionally, it can motivate pupils to speak and share their ideas and information. Moreover, they believe that it helps them raise their self-confidence, and improve their self-esteem. On the contrary, a few pupils still believe that group assessment can cause stress and test anxiety. Also, they reported that they will feel embarrassed if they make any mistakes. In general, these pupils cited many factors that would hinder their speaking performances.

Finally, regarding learners' attitudes towards the implementation of a Formal Oral Test, the results obtained from the final questionnaire and interview show that many of the pupils approved of the implementation of a Formal Oral Test. That is, they agreed on with the idea of arranging speaking tests in formal context for grades because of many reasons. First, they think that a Formal Oral Test will contribute to the improvement of pupils' speaking abilities. Second, it will develop pupils' self-confidence and self-esteem. Third, it will raise pupils' tests scores since many of them have good speaking abilities. Finally, the implementation of a Formal Oral Test will also enhance English language assessment and testing.

On the other hand, some pupils are against the idea because they believe that speaking tests are not necessary. They believe that pupils must focus only on reading and writing skills in order to succeed in their studies. In addition, they claim that stress, test

anxiety, increased pressure, and lack of confidence might negatively affect their speaking performance. Moreover, they argue that those pupils who do not speak English language well will get bad grades. Hence, the implementation of a Formal Oral Test will not be useful for them.

### **3.5 Synthesis of the Findings**

This section revisits the major issues, and tries to synthesise the findings which have been drawn from the analysis of the data obtained from the questionnaires and interviews. The rationale behind this is to answer the research questions and confirm the hypotheses of the current study.

First, the present study aims at exploring pupils' perceptions of English language learning and assessment. The results show that pupils are conscious of the fact that learning the English language requires mastering all its components. In other words, they are aware of the importance of each aspect of English language learning. They also know that they have to develop the four language skills. However, they believe that speaking is the most basic skill to be mastered due to its significance in daily life. Thus, since written tests do not assess their speaking abilities, they think that their test scores do not really reflect their levels in English language.

In addition, we attempted to investigate learners' attitudes towards the implementation of a Formal Oral Test, and to identify the factors that contribute to their attitudes. The findings drawn from the preliminary questionnaires in the first phase indicate that the pupils hesitated at the beginning. That is, they were not certain whether they had to agree with the idea of implementing a Formal Oral Test or not. However, the results obtained in the last phase of the study show that many of them think that it will be a good idea.

In terms of the factors associated with their positive attitudes, we believe that pupils formed their opinions based on their perceptions of English language learning and assessment. Besides, one of the factors that guided them to their opinions is the speaking test which they took in the second phase of this research. Many of them used words such as 'interesting', 'fun', 'happy', and 'excited' when describing the test. They also expressed their desire to take the speaking test again. Furthermore, there are other reasons which led to their positive attitudes towards the implementation of a Formal Oral Test and which were all discussed before. For instance, their beliefs that speaking tests will develop pupils' self-confidence and self-esteem, and it will improve pupils' test scores.

Moreover, this study sought to determine how learners perceive their experience of group speaking assessment. Similarly, if we compare the results of the first stage to those of the third stage, we will notice that more pupils now consider group assessment as the best way to assess pupils' speaking abilities. This is mainly because they recognised its advantages when they took the speaking test in the second stage of the research. For example, group format encourages pupils to interact with each other. It also provides them with a comfortable context for speaking assessment and this will, in turn, reduce stress levels and test pressure.

With reference to the present study findings, it can be deduced that group assessment had positive effects on the pupils' speaking performances, as well as their attitudes towards the incorporation of a Formal Oral Test. Therefore, we can say that all these results confirm the hypotheses that the current research is based on, and which are as follows:

- There is a relation between pupils' perceptions of English language learning and assessment and their attitudes towards the implementation of a Formal Oral Test;
- There is a positive relation between the use of group speaking assessment and pupils' speaking performances; and
- There is a positive relation between the use of group speaking assessment and pupils' attitudes towards the implementation of a formal oral test.

## **Conclusion**

To conclude, this final chapter has discussed the field work and data analysis. First of all, since research methodology is the basic element of any research, a theoretical background was provided to shed light on the methodology adopted to conduct the present study. Additionally, the chapter presented the different stages of this study. Then, it linked the findings to the research questions raised throughout the different phases in this investigation. We mainly attempted to prove that the implementation of a Formal Oral Test is a necessary step to reform English language assessment in Algeria. Besides, we sought to confirm the usefulness of group assessment as a way to assess pupils' speaking abilities.

Finally, it can be concluded that the chapter reported positive results.

## **General conclusion**

To reiterate, this study was based on the problem that the currently used tests in Algeria do not assess pupils' speaking abilities. Hence, they do not provide a reliable picture of pupils' achievement and progress. No one can deny that some changes have been introduced to the curriculum to enhance the level of middle school pupils in the English language. For instance, teachers are now encouraged to use different interaction patterns in the classroom, such as individual, pair and group work, in order to provide pupils with enough opportunities to speak in English language. Besides, the tutorial class is completely devoted to group tasks and oral communication activities. However, teachers still focus on teaching vocabulary and grammar, and ignore all these reforms. This is mainly because there are no considerable changes in formal assessment.

In an attempt to sensitise policy makers and test developers to the necessity for speaking tests, the present study aimed at investigating pupils' attitudes towards the implementation of Formal Oral Test. Additionally, we sought to examine the usefulness of group speaking assessment. In order to achieve the intended purposes, we adopted a qualitative research approach. Furthermore, we employed three data collection methods, namely a preliminary questionnaire, final questionnaire and interview, in order to collect the appropriate data. The findings of this research revealed that middle school pupils had positive attitudes towards the implementation of Formal Oral Test. Moreover, they showed interest in group format as it provided a comfortable context for speaking assessment.

With regard to the obtained results, we believe that this study shed light on the importance of speaking tests in order to achieve a comprehensive judgement on pupils' achievement and progress. Testing is not merely the operation of assigning scores. The currently used tests can be useful assessment tools for evaluating pupils in some specific areas. However, they do not assess pupils' speaking skills. Therefore, it can be argued that they do not yield reliable and valid scores.

Hopefully, this study has offered answers to the research questions. The research remains open to other further studies in the Algerian educational context.

## **Recommendations**

At last, some recommendations can be drawn from this study. These are as follows:

### **• For teachers**

- Teachers should give more attention to teaching the speaking skills.
- During all the sessions, teachers are required to use different interaction patterns to develop pupils' speaking abilities.
- Teachers should provide pupils with more opportunities to practise oral communication skills.
- Teachers should use authentic materials.
- Finally, teachers are required to be creative in order to provide motivating, and challenging learning environment.

### **• For pupils**

- Pupils are advised to practise speaking out of the classroom context in order to enhance their speaking abilities.
- Pupils are required to take advantage of opportunities offered in the classroom to practise their communication skills.
- It is important to be exposed to authentic language in order to develop vocabulary and acquire fluency.

### **• For administration**

- The administration should supply teachers with the necessary materials that can support teachers' efforts to develop pupils' speaking skills, such as computers, laptops and data show projectors.

### **• For researchers**

- Future researchers are recommended to conduct other studies in the language testing field. They are encouraged to consider speaking assessment in middle and high schools. Besides, they can carry out other researches to examine the effectiveness of the different speaking assessment formats, such as paired and group format.

### **• For policy makers and test developers**

- They are recommended to consider the implementation of a Formal Oral Test in order to develop pupils' speaking skills, as well as to improve English language learning and assessment. Additionally, they are recommended to consider the use of group speaking assessment for all the reasons discussed before.



## **References**

- Allison, D. (1999). *Language testing and evaluation: An introductory course*. Singapore: Singapore University Press.
- Bachman, L. F. (1990). *Fundamental consideration in language testing*. Oxford: Oxford University Press.
- Bachman, L. F. & Palmer, A. (1996). *Language testing in practise*. Oxford: University Press.
- Benjamin, A. (2013). *Formative assessment for English language art: A guide for middle and high school teacher*. New York: Routledge.
- Brassard, M. R. & Boehm, A. E. (2007). *Preschool assessment: Principles and practices*. New York: The Guilford Press.
- Brown, G. & Yule, G. (1983). *Teaching the spoken language: An approach based on the analysis of conversational English*. New York: Cambridge University Press.
- Brown, J. D. (2001). *Using surveys in language programs*. New York: Oxford University Press.
- Bryman, A. (2016). *Social research methods (5th Ed.)*. Oxford: Oxford University Press.
- Bygate, M. (1987). *Speaking*. Oxford: Oxford University Press.
- Carrasquillo, A. L. (1994). *Teaching English as a second language: A resource guide*. New York: Garland Publishing.
- Casas, M. (2011). *Enhancing student learning in middle school*. New York: Routledge.

- Cohen, L., Manion, L. & Morrison, K. (2007). *Research methods in education* (6th Ed.) London: Routledge.
- Comings, J., Garner, B. & Smith, C. (2006). *Review of adult learning and literacy, volume 6: Connecting research, policy, and practice*. New Jersey: Lawrence Erlbaum Associates.
- Creswell, J. W. (2014). *Research design: Qualitative, quantitative and mixed methods approaches* (4th Ed.). USA: Sage.
- Darling-Hammond, L. & Bransford, J. (2005). *Preparing teachers for a changing world: What teachers should learn and be able to do*. San Francisco: John Wiley & Sons.
- deMarrais, K. B. & Lapan, S. D. (2004). *Foundations for research: Methods of inquiry in education and the social sciences*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Denscombes, M. (2014). *The Good Research Guide: For small-scale social research projects* (5th Ed.). Maidenhead: Open University Press.
- Dörnyei, Z. (2007). *Research methods in applied linguistics: Quantitative, qualitative, and mixed methodologies*. Oxford: Oxford University Press.
- Elliott, J. (2005). *Using narrative in social research: Qualitative and quantitative approaches*. London: Sage.
- Flucher, G. (2014). *Testing second language speaking*. New York: Routledge.
- Gitsaki, C. & Coombe C. (2016). *Current issues in language evaluation, assessment and testing: Research and practice*. UK: Cambridge Scholars Publishing.
- Goh, C. C. M. & Burns, A. (2012). *Teaching speaking: A holistic approach*. New York: Cambridge University Press.

- Green, A. (2014). *Exploring language assessment and testing: Language in action*. New York: Routledge.
- Hoadjli, A. C. (2015). *The washback effect of an alternative testing model on teaching and learning: An exploratory study on EFL secondary classes in Biskra*. Biskra, Algeria: Mohamed Khider University of Biskra.
- Hoadjli, A. C. (2016). *Master 1 research methodology syllabus*. Biskra, Algeria: Mohamed Khider University of Biskra.
- Kothari, C. R. (2004). *Research methodology: Methods and techniques* (2nd Ed.). New Delhi: New Age International.
- Kumar, R. (2008). *Research methodology*. New Delhi: APH Publishing Corporation.
- Kumar, R. (2011). *Research methodology: A step-by-step guide for beginners* (3rd Ed.) London: Sage.
- Kunnan, A. J. (2000). *Fairness and validation in language assessment: Selected papers from the 19th Language Testing Research Colloquium, Orland, Florida*. Cambridge: Cambridge University Press.
- Kwon, M. (2015). *The effects of paired speaking assessment on Korean middle school students' speaking performance and their perceptions of speaking ability*. Montreal: McGill University.
- Long, M., Wood, C., Littleton, K, Passenger, T. & Sheehy, K. (2011). *The psychology of education* (2nd Ed.). Oxon: Routledge.
- Luoma, S. (2004). *Assessing speaking*. Cambridge: Cambridge University Press.
- McNamara, T. (2000). *Language testing*. Oxford: Oxford University Press.

- Merriam, S. B. & Tisdell, E. J. (2016). *Qualitative research: A guide to design and implementation*. 4th Ed. San Francisco: Jossey-Bass.
- Milanovic, M. & Weir, C. J. (2001). *European language testing in a global context: Proceedings of the ALTE Barcelona Conference, July 2001*. Cambridge: Cambridge University Press.
- Neuman, W. L. (2014). *Social research methods: Qualitative and quantitative approaches* (7th Ed.). Harlow: Pearson.
- Newman, I. & Benz, C. R. (1998). *Qualitative-quantitative research methodology: Exploring the interactive continuum*. USA: Southern Illinois University Press.
- Nunan, D. (1992). *Research methods in language learning*. New York: Cambridge University Press.
- Nunan, D. C. (1989). *Designing tasks for the communicative classroom*. Cambridge: Cambridge University Press.
- O'Reilly, K. (2012). *Ethnographic methods* (2nd Ed.). Oxon: Routledge.
- Orlich, D. C., Harder, R. J., Callahan, R. C., Trevisan, M. S, Brown, A. H. & Miller, D. E. (2013). *Teaching strategies: A guide to effective instruction* (10th Ed.). Belmont, USA: Wadsworth Cengage Learning.
- Punch, K. F. (2005). *Introduction to social research: Quantitative and qualitative approaches* (2nd Ed.). London: Sage.
- Richards, J. C. & Renandya, W. A. (2002). *Methodology in language teaching: An anthology of current practice*. New York: Cambridge University Press.
- Ryan, K., Cooper, J. M. & Tauer, S. (2013). *Teaching for student learning: Becoming a master teacher* (2nd Ed.). Belmont, USA: Wadsworth Cengage Learning.

Saraswathi, V. (2004). *English language teaching: Principles and practice*. Chennai, India: Orient Longman.

Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.

Thompson, S. K. (2012). *Sampling* (3rd Ed.). New Jersey. Wiley.

Wainer, H. & Braun, H. I. (2009). *Test validity*. Oxon: Routledge.

Watkins, D. C. & Gioia, D. (2015). *Mixed methods research*. New York: Oxford University Press.

## **APPENDICES**

## Appendix 1:

The Preliminary questionnaire: English version

### The Preliminary questionnaire

Dear pupils,

This research is conducted to explore fourth year pupils' perceptions of language learning and assessment in Cid Nouredine Middle School, Tolga. We would like to ask you for your opinions on this subject. Thank you for the time you are taking to complete this questionnaire. All answers will be held in the strictest confidentiality.

#### 1. Why do you study English?

- To satisfy school requirements.
- To be able to pass the Brevet Exam.
- To be able to read and write in English.
- To be able to communicate with foreigners.
- Other reasons

#### 2. In your view, what aspects of English language is the most important one?

- Vocabulary
- Grammar
- Pronunciation
- All the above

#### 3. Please rank the following language skills in order of importance from 1 to 4 where 1 is most important to you and 4 is least important to you.

- Listening
- Speaking
- Reading
- Writing

#### 4. Do achievement tests cover the content of what you have learnt?

- To a great extent
- To some extent
- To a very little extent
- No correspondence between the two.

**5. How are the testing methods used in these tests?**

- The same in every test.
- Vary slightly from one test to another.
- Vary completely from one test to another.

**6. What do you think the currently used tests assess?**

- Your degree of comprehension.
- The amount of vocabulary you know.
- Your mastery of grammatical structures.
- Your ability to communicate.
- All the above.

**7. How is the distribution of test-items among the test parts?**

- Equal
- Not equal
- It depends from one test to another.

**8. Which language skills are almost neglected in the currently used tests?**

- Listening
- Speaking
- Reading
- Writing

**9. Do you think the contents of these tests correspond with your interests?**

- Yes.
- No.
- Somehow.

**How?**

.....

.....

.....

.....

**10. Do your marks in tests usually reflect your progress and achievement in English language?**

- Yes.
- No.
- To some extent.



**11. To what extent you are able to express yourself through writing?**

- To a great extent.
- To some extent.
- To a very little extent.

**How?**

.....

.....

.....

.....

**12. To what extent you are able to express yourself through speaking?**

- To a great extent.
- To some extent.
- To a very little extent.

**How?**

.....

.....

.....

.....

**13. In your opinion, should listening and speaking skills be given more importance in English language tests?**

- Yes.
- No.

**Why?**

.....

.....

.....

.....

**14. Do you think that the implementation of a Formal Oral Test will be a good idea to improve English language assessment?**

- Yes.
- No.

**Why?**

.....

.....

.....

.....

**15. How would you like your speaking performance to be assessed?**

- Individually.

- In pairs.

- In groups.

**Why?**

.....

.....

.....

.....

**End of Questionnaire**

**Thank you very much for your help**

## Appendix 2:

### The Preliminary questionnaire: Arabic version

#### الإستبيان الأولي

أعزائي التلاميذ،  
هذا البحث يهدف إلى معرفة تصورات تلاميذ السنة الرابعة بمتوسط الصيد نور الدين بطولقة حول تعلم اللغة الانجليزية و عملية التقييم. ولهذا نرغب بمعرفة آرائكم بخصوص هذا الموضوع. نتقدم بالشكر الجزيل على الوقت الذي ستمنحونه لإستكمال هذا الاستبيان. ونود ان نعلمكم ان أجوبتكم سوف تدرس بسرية تامة.

1. لماذا تدرس اللغة الانجليزية؟

تلبية متطلبات المدرس

لاجتياز امتحان شهادة التعليم المتوسط

لكي تكون قادر على التواصل مع الأجانب

لأسباب أخرى

2. حسب رأيك، اي جانب من جوانب اللغة الانجليزية هو الأكثر أهمية؟

مفردات اللغة

قواعد اللغة (النحو)

علم الاصوات ( النطق )

كل ما سبق

3. من فضلك قم بترتيب المهارات اللغوية التالية من حيث الاهمية من 1 الى 4، حيث 1 هو الاكثر أهمية بالنسبة لك و 4 هو الأقل أهمية بالنسبة لك.

الاستماع

التكلم

القراءة

الكتابة

4. هل الإختبارات التحصيلية تغطي محتوى ما تعلمته؟

الى حد كبير

الى حد ما

الى حد ضئيل

ليس هناك علاقة بينهما على الاطلاق

5. كيف هي الأساليب والطرق المستخدمة لبناء الاختبارات؟

متشابهة

تختلف نسبيا من إختبار الى آخر

تختلف كلياً من إختبار الى آخر

6. حسب رأيك، ماذا تقيم الإختبارات المستخدمة الحالية؟

مدى فهمك

كم المفردات التي تعرفها

مدى تمكنك من التراكيب النحوية

قدرتك على التواصل

كل ما سبق

7. كيف ترى توزيع عناصر الاختبار على مختلف اقسامه؟

متساوي

غير متساوي

يختلف من اختبار لآخر

8. ما هي المهارة اللغوية المهمة إلى حد ما في الاختبارات المستخدمة حالياً؟

الاستماع

التكلم

القراءة

الكتابة

9. هل تعتقد أن محتويات الاختبارات الحالية تتوافق مع اهتماماتك؟

نعم

لا

إلى حد ما

كيف ذلك؟

.....

.....

.....

.....

10. هل علامتك في الإختبارات تعكس تقدمك و إجازاتك في اللغة الانجليزية؟

نعم

لا

إلى حد ما

11. إلى أي مدى تستطيع التعبير عن نفسك عن طريق الكتابة؟

إلى حد كبير

إلى حد ما

إلى حد ضئيل

كيف ذلك؟

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

12. إلى أي مدى تستطيع التعبير عن نفسك طريق الكلام؟

إلى حد كبير

إلى حد ما

إلى حد ضئيل

كيف ذلك؟

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

13. في رأيك، هل يجب اعطاء كل من مهارتي الاستماع و التكلم أهمية أكبر في اختبارات اللغة الانجليزية؟

نعم

لا

لماذا؟

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

14. هل تظن أن فكرة ادراج اختبار شفوي رسمي فكرة جيدة لتحسين تقييم اللغة الانجليزية؟

نعم

لا

لماذا؟

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

15. كيف تفضل ان يكون تقييم اداءك الشفهي؟

فردي

زوجي

جماعي

لماذا؟

.....

.....

.....

.....

نهاية الاستبيان

شكرا جزيلا لتعاونكم

### Appendix 3:

The Final questionnaire: English version

#### The Final questionnaire

Dear pupils,

This research is conducted to explore fourth year pupils' perceptions of English language learning and assessment, and their attitudes towards the implementation of a Formal Oral Test. We would like to ask you for your opinions on this subject. Thank you for the time you are taking to complete this questionnaire. All answers will be held in the strictest confidentiality.

#### 1. How was the speaking test?

- Very easy
- Easy
- Neutral
- Difficult
- Very difficult

#### 2. Were the questions/tasks the ones you had never encountered or expected?

- Yes.
- No.
- Somehow.

#### 3. Were the questions/tasks in the speaking test explicit and clear?

- Yes.
- No.
- Somehow.

#### 4. How did you feel during the speaking test?

.....

.....

.....

.....

#### 5. Did the assessor encourage and motivate you during the speaking test?

- Yes.
- No.
- Somehow.

**6. What were the difficulties that you faced during the test?**

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

**7. What did the speaking test allow you to do?**

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

**8. In your view, does group format provide an effective and a comfortable context for speaking assessment?**

- Yes.
- No.
- Somehow.

**How?**

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

**9. Please rank the following language skills in order of importance from 1 to 4 where 1 is most important to you and 4 is least important to you.**

- Listening
- Speaking
- Reading
- Writing

**10. In your view, what aspect of English language is the most important one?**

- Vocabulary
- Grammar
- Pronunciation
- All the above



**11. Now, do you think that your marks in currently used tests usually reflect your progress and achievement in English language?**

- Yes.
- No.
- To some extent.

**Why?**

.....

.....

.....

.....

**12. Do you think that speaking test should be administered in a formal context for grades?**

- Yes.
- No.

**Why?**

.....

.....

.....

.....

**End of Questionnaire**

**Thank you very much for your help**

## Appendix 4:

### The Final questionnaire: Arabic version

#### الإستبيان النهائي

أعزائي التلاميذ،

هذا البحث يهدف إلى معرفة تصورات تلاميذ السنة الرابعة بمتوسط الصيد نور الدين بطولقة حول تعلم اللغة الانجليزية و عملية التقييم، و كذلك معرفة مواقفهم من فكرة إدراج إمتحان شفهي رسمي. ولهذا نرغب بمعرفة آرائكم بخصوص هذا الموضوع. نتقدم بالشكر الجزيل على الوقت الذي ستمنحونه لإستكمال هذا الاستبيان. ونود ان نعلمكم ان أجوبتكم سوف تدرس بسرية تامة.

1. كيف كان الإمتحان الشفهي؟

- سهل جدا

- سهل

- حيادي

-صعب

- صعب جدا

2. هل كانت الاسئلة / المهام تلك التي لم تختبرها من قبل او لم تتوقعها؟

- نعم

- لا

- الى حد ما

3. هل كانت أسئلة / مهام الإختبار الشفهي مباشرة وواضحة؟

- نعم

- لا

- الى حد ما

4. كيف شعرت خلال الإختبار الشفهي؟

.....

.....

.....

.....

5. هل قام الأستاذ المقيم بتشجيعك و تحفيزك أثناء الإختبار الشفهي؟

- نعم

- لا

- الى حد ما

6. ماهي الصعوبات التي واجهتها خلال الإختبار الشفهي؟

.....

.....

.....

.....

7. ماذا سمح أو أتاح لك الإختبار الشفهي أن تفعل؟

.....

.....

.....

.....

8. في رأيك، هل يوفر التقييم في جماعات السياق المريح و الفعال لعملية تقييم مهارة الكلام؟

نعم

- لا

- الى حد ما

كيف ذلك؟

.....

.....

.....

.....

9. من فضلك قم بترتيب المهارات اللغوية التالية من حيث الأهمية ، حيث 1 هو الأكثر أهمية بالنسبة لك و 4 هو الأقل أهمية بالنسبة لك.

-الإستماع

- التكلم

- القراءة

- الكتابة

10 . حسب رأيك، أي جانب من جوانب اللغة الإنجليزية هو الأكثر أهمية؟

- مفردات اللغة

- قواعد اللغة (النحو)

- علم الأصوات (النطق)

- كل ما سبق

11. الآن، هل تعتقد ان علامتك في الإختبارات المستخدمة حاليا عادة ما تعكس تقدمك وإنجازاتك في اللغة الإنجليزية؟

نعم

- لا

- الى حد ما

كيف ذلك؟

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

12. هل تعتقد انه يجب إعتداد إختبار شفهي بشكل رسمي من أجل العلامات؟

نعم

- لا

لماذا؟

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

نهاية الاستبيان

شكرا جزيلاً لتعاونكم

## **Appendix 5:**

### **The Interview**

#### The Questions:

Q1: How long have you been studying English?

Q2: Do you like English? Why?

Q3: There are four language skills which are listening, speaking, reading and writing. In your opinion, which language skill is the most important one? Why?

Q4: You took an English speaking test a couple of days ago. Is that right? How was the test? How can you describe that speaking test?

Q5: How did you feel during the speaking test?

Q6: What were the difficulties that you faced during the test?

Q7: What did the speaking test allow you to do?

Q8: In your opinion, does group format provide an effective and a comfortable context for speaking assessment?

How is that?

Q9: Do you think that speaking test should be administered in a formal context for grades? Why?

-End of Interview-

Thank you for your collaboration.

**Appendix 6:**  
**The speaking test**

**English Speaking Test**

Good morning. How are you today?

**Part 01**

First, I would like to know something about you.

Q1: Who do you spend time with after school?

Q2: Do you enjoy reading books? Why?

Q3: Do you enjoy using the Internet in your free time?

Q4: Where would you like to go for your next holiday?

**Part 02**

In this part of the test I'm going to give each one of you a photograph. I'd like you to talk about your photographs on your own for about a minute.

Here are the photographs. They show people spending time together in different situations.



The Questions:

What are the people enjoying about spending time in these situations?

Which one of these would you prefer to do? Why?

### **Part 03**

Now, I would like you to talk about something together for about two minutes.

I'd like to imagine that the school is going to start some after school classes to encourage the pupils to learn new skills. Here are some ideas for the classes and a question for you to discuss. First, you have some time to look at the tasks.

**Speak another foreign language**

**Play a musical  
instrument**

**Why might pupils want to  
learn to do these skills?**

**Draw well**

**Cook a meal**

**Play a sport**

Now, talk to each other about why pupils might want to learn these skills.

The Questions:

Do you think classes like these would be popular with pupils? Why?

How important do you think it is for people to try new activities?

Thank you. That's the end of the test.

## Glossary

**Assessment:** The process of quantifying the characteristics of persons according to explicit procedures and rules.

**Evaluation:** It is the gathering of pertinent and reliable information, such as verbal descriptions and comprehensive impressions in order to evaluate learners.

**Formative assessment:** It is the type of assessment that is used to monitor learning progress during instruction.

**Speaking:** It is an integral part of people's daily lives. It can be defined as the process of encoding the message the speaker intends to convey using appropriate language.

**Summative assessment :** It is the type of assessment that is developed to measure students' achievement using grades.

**Test:** It is one type of measurement tools which is devised to obtain a representative sample of an individual's behaviour.



## ملخص

تمحورت هذه الدراسة حول موضوع عملية التقييم لمادة اللغة الإنجليزية في الجزائر، وبصيغة أدق حول الإختبارات التحصيلية لمادة اللغة الإنجليزية في الطور المتوسط.

تهدف هذه الدراسة الى تحديد مواقف التلاميذ من فكرة إدراج اختبار شفهي رسمي لمادة اللغة الإنجليزية. بالإضافة الى تحديد تصوراتهم حول عمليتي التعلم والتقييم. الى جانب ذلك، تسعى الى تحديد مدى فعالية تقييم مهارات التلاميذ الكلامية في جماعات.

وبناء على هذا فإن الدراسة الحالية تتكون من جزأين وهما الجانب النظري، الذي يسلط الضوء على عملية التقييم و مهارة الكلام، والجانب التطبيقي، والمتمثل في الدراسة الميدانية و تحليل المعطيات .

اعتمدنا في عملية جمع المعطيات على استعمال استبيانين و مقابلة شفوية مع تلاميذ السنة الرابعة متوسط بمتوسطة الصيد نور الدين بطولقة.

في الأخير يمكن القول ان النتائج كانت ايجابية حيث أن أغلبية التلاميذ أيدوا فكرة ادراج اختبار شفهي رسمي في الطور المتوسط. كما انهم استحسنا فكرة تقييم مهارات التلاميذ الكلامية في جماعات لانها توفر سياق مريح و تتيح التفاعل بين التلاميذ.